IDENTIFYING PROTEIN-PROTEIN INTERACTION INHIBITORS TARGETING APICOMPLEXAN ATG8-ATG3 INTERACTIONS

by
Alexia S. Miller

A dissertation submitted to Johns Hopkins University in conformity with
the requirements for the degree of Doctor of Philosophy

Baltimore, MD
July 2016

ABSTRACT

Malaria is a global health crisis that threatens almost half of the world's population. The need for novel antimalarials is becoming increasingly urgent as the incidence of resistance rises against the first-line defense, artemisinin. *Plasmodium falciparum*, the causative agent of the most lethal form of malaria, is a member of the phylum Apicomplexa. This phylum of parasitic protists is responsible for a number of diseases afflicting both human and non-human hosts, resulting in significant health and economic burdens worldwide. Recent work has pointed to the conserved autophagy pathway within these parasites as a potential target for anti-apicomplexan drug development.

In this body of work, two approaches are taken to identify novel small molecules targeting the autophagy pathways within *Plasmodium*, and within other apicomplexans, as well.

The first approach involves a Virtual Ligand Screen of the ChemBridge™ Express Library against the crystal structure of *P. falciparum* Atg8. Hits are ordered and tested using SPR to detect activity against the plasmodial interaction without targeting the human interaction. The top hit, ALC25, is verified *in vitro* using SPR dose dependency, thermal shift, and ITC; it is additionally confirmed *in vivo* against both blood- and liver-stage *P. falciparum* parasites. ALC25 exhibits modest antimalarial activity, with IC_{50} values below 20 μM.

The second approach expands upon hits identified from the Medicines for Malaria Venture (MMV) Malaria Box. These hits were detected via a medium-throughput Surface Plasmon Resonance (SPR) interaction assay and were subsequently verified *in vivo*. The small molecule scaffold for these hits is then chemically diversified and tested, again using SPR, against the Atg8-Atg3 interactions from *Plasmodium falciparum*, *Cryptosporidium parvum*, *Eimeria tenella*, and *Neospora caninum*/*T.gondii*, with promising results.

These compounds, targeting the Atg8 conjugation pathway within *P. falciparum*, are then used to begin exploring the novel connection between the plasmodial autophagy pathway and the export of *P. falciparum* Erythrocitic Membrane Protein 1 (*Pf*EMP1), a connection with profound implications for the potential treatment of cerebral malaria.

First reader and thesis advisor: Dr. Jürgen Bosch

Second reader: Dr. Sean Prigge

ACKNOWLEDGEMENTS

I would like to acknowledge all of those individuals who have made this journey possible. First and foremost, I express my heartfelt gratitude toward Jürgen Bosch and James Berger as my research advisors.

Thank you, Jürgen, for your enthusiastic, problem-solving, silver-lining-finding attitude. And thank you for continually encouraging me to test my limits! I learned so much from you, not only about science, but about how to be passionate about the things that matter.

To James, I am inexpressibly grateful for the help you extended in a time of need and for giving me the incredible opportunity to learn from you and the brilliant members of your lab.

Thank you to my committee members – Cynthia Wolberger, Sean Prigge, Scott Bailey, and Herschel Wade – for your consistently constructive feedback and for taking an interest in my wellbeing in graduate school and beyond.

To my classmates, I am inordinately fond of all of you. You could make even the most difficult periods of graduate school into the best memories.

To all labmates past and present, thank you for your expert advice, collaboration, and friendship. Ulli, I am particularly indebted to you for mentoring me into my first graduate school lab and for your collaboration and help all along the winding road to getting published.

And to the members of the Berger lab, thank you for your acceptance, assistance, and amiability throughout. You are all amazing!

I am also profoundly grateful to my family, without whom none of this would have been possible. Thank you for your belief in me, for your words of encouragement, for your love, for your editing, and for simply making the world a better place by the mere fact of your existence.

Finally, to my husband Geoffrey, words cannot express what it means to me to have your unshakable, unconditional love and support. I love you to the stars and beyond. And to my daughter Diella, the light of my life, thank you for inspiring me to be a better person every day. You fill my life and my heart with joy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction to Targeting Autophagy in Apicomplexa

Sections published in:

Miller, A.S., and Bosch, J. (2016) Targeting the Atg8 Conjugation Pathway for Novel Anti-Apicomplexan Drug Discovery, in *Comprehensive Analysis of Parasite Biology: From Metabolism to Drug Discovery*, 1ed., Wiley-VCH, Weinheim. (Reprinted with permission)

Abstract

Autophagy is an important metabolic pathway, conserved within eukaryotes, for the degradation of superfluous or damaged proteins and organelles. In addition to its role in homeostasis and survival under starvation conditions, autophagy has been found to function in a variety of pathways including cell differentiation and development, cell-mediated death, and clearance of intracellular pathogens. Within the phylum Apicomplexa, however, autophagy has increasingly been found to serve novel roles specific to intracellular parasite biology. Apicomplexa are obligate intracellular parasitic protists that represent a significant global burden to human health and economic development. Autophagy within these parasites represents an attractive target for novel anti-apicomplexan therapies. Here, recent advances in the field of apicomplexan autophagy research are presented. We also discuss recent results from our laboratory, in which we use two different approaches to identify novel lead-like molecules that inhibit the *Pf*Atg3-*Pf*Atg8 protein–protein interaction. These compounds have activity against blood and liver stage *Plasmodium falciparum* malaria and lead to an upregulation and shift in lipidation state of *Pf*Atg8. The molecules and methods identified from these two complementary approaches have applications for apicomplexan small-molecule intervention strategies in general.

Relevance and introduction to thesis

For a disease like malaria, eradication cannot be realized from a single approach. Recent years have seen a mounting emphasis on prevention, in the form of insecticide-treated nets and indoor residual spraying [1], and vaccine development [2]. These approaches offer significant appeal. Anti-vector techniques alone are estimated to have prevented more than 500 million cases in the period between 2000 and 2015 [3], and the RTS,S/AS01 malaria vaccine has been shown to have some limited efficacy in the broader population: 36 percent vaccine efficacy against clinical malaria when administered in three doses with an additional booster [4]. However, the RTS,S/AS01 malaria vaccine presents no significant protection against severe malaria; in fact, without the booster dose, it may contribute to cases of severe malaria by preventing the natural progression of immunity [4].

More generally, neither preventative measures nor current vaccines offer solutions for the millions of people who will become sick every year at the hands of this disease. Not only does this represent a substantial burden of human suffering, but it also influences the achievability of any plan for malaria elimination: these individuals represent an ever-present pool of disease carriers. As long as prevention and vaccination remain imperfect, malaria will continue to wreak havoc on the human population in the absence of effective medication. Thus, for the sake of human health in the moment and eradication in the long run, drug development must also play a role.

For those who suffer from this illness, current treatment options are frighteningly limited. With chloroquine resistance nearly universal, artemisinin-based combination therapies (ACTs) are the go to treatment. Now, however, even ACTs have been threatened as reliable treatment options: resistance has been detected in at least five countries and is spreading over time [5]. As we deal with mounting resistance to ACT, new antimalarials are desperately needed.

This urgent need forms the foundation for this body of work. My doctoral research focuses on the search for antimalarials targeting the *Plasmodium falciparum* autophagy system. In the first chapter, I discuss the current understanding of autophagy within the phylum Apicomplexa, of which *Plasmodium* is a member, and briefly discuss two approaches we have taken to target the plasmodial autophagy system. In the second chapter, I detail a Virtual Ligand Screening approach to drug discovery. In the third chapter, I present the results of further testing of the PTA derivative molecules previously found to inhibit an important autophagic protein-protein interaction in *P. falciparum*. Finally, chapter four concludes with the importance of this work and future directions for this research.

Autophagy: an overview

Macroautophagy, or more simply autophagy, is a catabolic process within eukaryotic cells responsible for the recycling of intracellular material. It is mediated by double-membrane-bound vesicles called autophagosomes, which engulf cytoplasmic material and organelles before fusing with the lysosome to form autolysosomes [6]. This process is widely conserved in eukaryotes [7] and is involved in a variety of functions within the cell. Autophagy facilitates cellular homeostasis, occurring constitutively at a basal level to regularly turn over damaged or unused cellular debris and organelles [8,9]. Under amino acid starvation conditions, autophagy is significantly upregulated in order to internally scavenge the materials necessary for continued survival [10]. In addition to functioning in homeostasis and the stress response [11,12], autophagy is also involved in a variety of other cellular processes. For example, autophagy serves an important role in cell differentiation and development, where it facilitates rapid changes in the cellular environment through degradation of unnecessary intracellular structures and organelles [13]. As a natural extension of its role in engulfing unwanted intracellular components, autophagy is also involved in the clearance of intracellular pathogens through a process termed xenophagy [14]. Finally, in a seeming contradiction to its role in cell survival under amino acid starvation conditions, autophagy also appears to be involved in non-apoptotic programmed cell death under specific cellular conditions [15].

This complex range of functions is performed by a set of more than 30 autophagy (Atg) proteins that have been well-characterized in yeast [16]. It is useful to have an overview of the main proteins and complexes involved in performing macroautophagy within yeast before taking a closer look at autophagy within Apicomplexa [17] (Figure 1.1A). Initiation, the first step in autophagy, begins when the Atg1 initiation complex forms at the Pre-Autophagosomal Structure (PAS), the site from which the autophagosome originates [18]. The formation of the Atg1 complex is controlled by negative regulation through Target Of Rapamycin Complex 1 (TORC1), the inhibitory function of which can be blocked by exposure to amino acid starvation conditions or treatment with rapamycin [19]. The Atg1 complex consists of the kinases Atg1 and Atg13, which bind to the ternary complex formed by Atg17 bound to Atg31 and Atg29 [20]. Once formed, the Atg1 complex activates the class III phosphatidyl-inositol 3-phosphate kinase (PI3K) complex (comprising: Vps34, Vps15, Vps30/Atg6, Atg14, and Atg38) [21,22]. The activated PI3K class III complex initiates phagophore assembly, converting phosphatidylinositol (PI) into phosphatidylinositol triphosphate (PI3P) at the PAS. Wortmannin and 3-methyladenine, classic autophagy inhibitors, inhibit PI3K activity [23]. The existence of PI3P at the PAS then recruits several autophagy-related proteins (Atg2, Atg9, and Atg18) to bind to the incipient phagophore [24]. Although its role remains somewhat unclear, Atg9 is thought to transport membrane to the phagophore at this stage [25]. Elongation and closure of the phagophore, to form a completed autophagosome, require conjugation of Atg8's C-terminal exposed glycine residue to phosphatidylethanolamine (PE) within the phagophore membrane [26].

Atg8 is a ubiquitin-like protein (Ubl) whose conjugation pathway involves the enzymatic function of several Atg proteins. Atg4 is a protease that cleaves the C-terminal tail of Atg8, exposing a conserved glycine residue necessary for Atg8 conjugation with PE; Atg4 also functions to deconjugate Atg8 from the autophagosomal membrane [27]. Atg7 is an E1-activating enzyme; Atg3 is the E2-conjugating enzyme actually responsible for conjugating Atg8 to PE via Atg8's exposed C-terminal glycine; and the covalently bound Atg12–Atg5 proteins form a complex with Atg16 to significantly increase Atg8 conjugation, by acting as an E3-like enzyme [28–30]. Atg12 is also a Ubl, requiring the same E1 enzyme, Atg7, and its own E2 enzyme, Atg10, to covalently attach it to Atg5 [28]. Once Atg8 has been conjugated to the membrane, the phagophore elongates, engulfing the intracellular material, and then closes before fusing with the lysosome.

Autophagy is increasingly appreciated for its roles, not only within healthy cells but also within diseased or infected cells, where, in many cases, it is shown to have both pro- and anti-survival effects, depending on the specific cellular environment. For all of these reasons, research into this process is essential for developing our understanding of fundamental cell function and disease progression. However, while this pathway continues to be well-explored within yeast and other higher eukaryotes, the situation is much less clear for protists, including those in the phylum Apicomplexa.

Phylum, Apicomplexa

Apicomplexa is a phylum of parasitic protists that represent a significant threat to human health and economic development. *Plasmodium*, the best known member of the phylum, is responsible for the disease malaria, which kills about 500,000 people every year (www.who.int/topics/malaria). While *Plasmodium* is undoubtedly the deadliest organism in the phylum, many other Apicomplexa present a significant threat to immunocompromised individuals or to the livestock industry. *Toxoplasma gondii*, the causative agent of toxoplasmosis, infects nearly one third of the world's population [31]. While *T. gondii* infection is predominantly harmless in healthy individuals, toxoplasmosis can be deadly for HIV patients and pregnant women, whose unborn children can acquire severe health complications through congenital infection (www.cdc.gov/toxoplasmosis). *Cryptosporidium parvum* is another organism that threatens the immunocompromised through the disease cryptosporidiosis, one of the major causes of waterborne illness globally [32]. For the agricultural industry, there are several organisms of considerable concern. *Neospora caninum*, *Babesia bovis*, and *Theileria annulata* all affect the cattle industry; *N. caninum* is one of the primary causes of fetal death in cows [33]. *Eimeria tenella* afflicts the poultry industry, causing caecal coccidiosis in young poultry [34]. It is difficult to estimate the cost of these various infections to the livestock industry, though *N. caninum* alone is estimated to lead to losses of over half a billion dollars annually to the U.S. cattle industry [35]. Because of these economic losses, apicomplexan infections are not only a considerable threat to human health and wellbeing, but also a threat to the growth of developing nations due to the sheer burden of these diseases on a growing economy.

While various anti-apicomplexan therapies are in use, not all are wholly effective or cost-effective. Of pressing concern is the need for novel antimalarials, as increasing reports of drug resistance have been emerging against artemisinin, the first line of defense [5,36]. To develop new classes of anti-apicomplexan therapies, it is essential to identify viable targets within these eukaryotic pathogens, targets with enough divergence from host systems to ensure specificity and limit off-target effects.

In terms of their basic biology, Apicomplexa are characterized by several shared features. They are obligate intracellular parasites, with complex life cycles spanning multiple hosts. They are classified based on a conserved “apical complex,” a structure used for host cell invasion with specialized secretory organelles, including micronemes and rhoptries [37]. Most of the organisms (*C. parvum* is a notable exception) feature a unique plastid organelle called the apicoplast, which was acquired through secondary endosymbiosis, resulting in a structure bound by four membranes [38]. This organelle functions in the synthesis of fatty acids, isoprenoids, iron-sulfur clusters, and heme [38]. Added to the shared physical characteristics of Apicomplexa is the additional common feature of spending the majority of their life cycle within host cells. Because of this distinct cellular environment, Apicomplexa differ from other unicellular eukaryotes in a myriad of ways; this extends to the process of autophagy, as well.

Conservation of autophagy in Apicomplexa

Autophagy within Apicomplexa represents an important and growing area of research, not only as a method to better understand the biology of these parasites, but also as a potential target for future drug discovery efforts. At this stage, the genomes for several parasites within this phylum are available, including those species listed above [39]. It is important to remember the level of diversity represented within a phylogenetic group, a point which is of particular relevance for Apicomplexa, as they have considerable intra-phylum genetic variation [40]. In spite of this level of diversity, it is still possible to obtain a general sense for the degree of conservation of the autophagy proteins in the phylum. Based on a variety of genome analyses, homologues for a number of proteins have been identified [17,41–49]. The *P. falciparum* genome, in particular, has been subjected to numerous searches for Atg homologues. As the majority of homologues have been determined via computational rather than experimental analyses, there is a considerable amount of disagreement in the results of these investigations. From the aforementioned screens, various *P. falciparum* homologues have been proposed for the following: Atg1, Atg2, Atg3, Atg4, Atg5, Vps30/Atg6, Atg7, Atg8, Atg10, Atg12, Atg14, Atg15, Vps15, Atg16, Atg17, Atg18, Atg22, Vps34, and TOR [17,41–47,49]. For any single analysis, many fewer homologues were identified (e.g., Figure 1.1B). Similarly, for *T. gondii*, Atg1, Atg3, Atg4, Atg5, Vps30/Atg6, Atg7, Atg8, Atg9, Atg12, Vps15, Atg17, Atg18, Vps34, and TOR homologues have been presented [48,49]. Other Apicomplexa, including *N. caninum*, *E. tenella*, *C. parvum*, *C. hominis*, and *Theileria* spp. have also been investigated [41–43,48,49].

Given the range of results that have been reported, it is therefore noteworthy that the entire Atg8 conjugation pathway (Atg3, Atg4, Atg7, and Atg8) has consistently been found to be conserved in Apicomplexa. For this reason, the identified homologues have received special focus. From a number of studies, it appears that the pathway is not only conserved, but also essential to parasite survival. The crucial nature of the Atg8 conjugation pathway represents a point of divergence from other eukaryotes, as it has been shown that many Atg protein knockout strains in yeast or mammalian organisms are still viable [13,26]. This essentiality has been demonstrated for three Atg proteins in *T. gondii*: interference with the Ubl *Tg*Atg8 leads to non-viable parasites [50]; elimination of the protease *Tg*Atg4 results in significant growth defects [50]; and the E2-like enzyme *Tg*Atg3 is also essential for normal organelle morphology and intracellular development [48]. Similarly, in *Plasmodium*, *Pf*Atg8 knockouts are nonviable [42,43] and *Pf*Atg7 is essential for normal growth [47].

Functions of apicomplexan autophagy proteins

From these studies, it is clear that Apicomplexa make use of, and require, a much more limited repertoire of autophagy proteins. But to what end? Are these organisms performing canonical autophagy, or is this more limited set of autophagy proteins serving different purposes altogether? A growing body of research has arisen to provide some answers to these questions. In particular, more information about the function of the Atg proteins within *Toxoplasma* and *Plasmodium* is now available, though significant questions still remain to be answered.

Atg8 in yeast and other eukaryotes is synthesized with a C-terminal extension that must be cleaved to expose the glycine residue necessary for conjugation to PE [51,52]. In Apicomplexa, however, Atg8 homologues contain a glycine as the C-terminal residue, implying that the protein does not need to be modified by proteolytic cleavage prior to conjugation to PE and, presumably, a membrane. This is true for potential homologues of Atg8 identified in *B. bovis* (EuPathDB gene id: BBOV_III005950), *E. tenella* (ETH_00016760), *N. caninum* (NCLIV_008410), *Plasmodium* spp. (e.g., *PF*3D7_1019900), *T. annulata* (TA12615), and *T. gondii* (*TG*ME49_254120) [39]. *C. parvum* Atg8 (cgd7_3990) is an exception [39]. Perhaps because it is already in this active form, Atg8 in both *Plasmodium* and *Toxoplasma* exists in a predominantly lipidated, membrane-associated form [46,53,54]. Although Atg8 does not need to be pre-processed by Atg4 in these parasites, Atg4 is nevertheless essential to intracellular development in *T. gondii*, and a role in Atg8 recycling from the membrane was proposed, which appears to be essential for the parasites [27,50].

The observation that the majority of Atg8 in these organisms is membrane-bound leads to one of the most unequivocally established connections to autophagy made within Apicomplexa: in numerous studies on *T. gondii* and *Plasmodium* species, Atg8 has been found to partially or completely co-localize with the apicoplast membrane [55]. Its role in this location is unclear, though it appears that Atg8 is conjugated to PE in the outermost membrane of the organelle [45,50,56], and it is not engulfing the apicoplast within an autophagosome [45,53,54]. For these reasons, and because ablation of Atg4 function leads to significant defects in apicoplast morphology [50], Atg8 is thought to play a fundamental role in the biogenesis and maintenance of the apicoplast [45]. However, this is not the only function of the protein within the parasite.

Apicomplexa undergo significant morphological changes over the course of their life cycle, changes which must be accompanied by rapid remodeling of the cellular environment with the creation and disposal of stage-specific organelles. As mentioned in the overview of autophagy in higher eukaryotes, autophagy plays a pivotal role in mammalian cell differentiation and development [13,57]; these results also apply to certain eukaryotic pathogens, including pathogenic strains of the fungus *Candida* [58,59]. More significantly, autophagy has also been implicated in the cellular differentiation and homeostasis of trypanosomatids [60], particularly *Trypanosoma brucei* [61], *Trypanosoma cruzi* [62,63], and *Leishmania* spp. [64–67], organisms which undergo similarly complex changes in morphology to adapt to diverse host cell environments. For these reasons, a role for autophagy in apicomplexan development and differentiation has been explored. There are clear indications of this in *T. gondii*, since, as mentioned above, *Tg*Atg3 is essential to normal cellular development: loss of *TgAtg3* led to significant morphological and growth defects. [48]. For *P. falciparum*, the parasite enters the human host as an elongated sporozoite, a highly motile and invasion-competent form of the parasite that penetrates a hepatocyte before transforming into a round, metabolically active trophozoite that then replicates [68]. This metamorphosis is accompanied by reorganization of cytoskeletal structure and clearance of the rhoptries, micronemes, and inner membrane complex [69]. Autophagy has been proposed as one mechanism for this clearance. In support of this proposal, membranous structures enclosing micronemes have been detected via electron microscopy during this metamorphosis [69], and 3-MA delays conversion of sporozoites to trophozoites [70].

Further exploring this idea of autophagosomes being coopted for exophagy, there is some evidence to suggest that the autophagy pathway within Apicomplexa may serve a role in trafficking to the host cell. Both *Pf*Atg8 and *Pf*PI3K have been found near the Maurer's cleft [71,72], a structure essential to vesicular traffic to the red blood cell (RBC) membrane [73]. This idea may be further supported by the existence of *Pf*Atg8 punctate structures throughout the RBC [46], as these structures could be involved in the export necessary for extensive host cell remodeling, although these structures have also been proposed to serve other functions. However, the existence of Atg8 in these punctate structures and near the Maurer's cleft remains a disputed finding, as several studies have found no evidence of *Pf*Atg8 aside from its apicoplast association [53,54,56]. These contradictory results may arise from differences in protein constructs and detection methods. Either way, further investigation is clearly necessary to shed light on whether or not

*Pf*Atg8 is in fact exported to the host cell, and, if so, what its function may be there.

The apicomplexan response to amino acid starvation is an area of some debate, as various studies have pointed to parasites engaging in a variety of responses, including hibernation, autophagosomal scavenging, or autophagic programmed cell death. This is not unexpected: as mentioned earlier for other eukaryotes, autophagy often plays a role in both pro-survival and pro-death functions. Indications of a pro-survival role arise in response to starvation conditions within the cell. In particular, nutrient or amino acid deprivation in both *T. gondii* [48,74] and *P. falciparum* [45,75] parasites led to *Pf*Atg8 punctate structures throughout the cell. In *P. falciparum*, *Pf*Atg8 was found to co-localize with *Pf*RAB7 upon amino acid starvation, the plasmodial homologue of RAB7, an important protein in autolysosomal formation in higher eukaryotes [45]. Notably, these proteins were also detected in association with the food vacuole [45]. However, not every study has found the same response to starvation: in *P. falciparum*, hibernation without a detectable upregulation of Atg8 production was also observed [46,76]. Furthermore, the identities of the various punctate structures have not been unequivocally determined, so it remains to be determined whether or not these structures correspond to classical autophagosomes.

On the other side of the life/death balance, there is evidence for autophagic programmed cell death in response to various signals, including nutrient deprivation, drug treatment, heat stress, and even normal cellular events. For example, extensive vacuolization has been observed in dying liver stage parasites in mice, though *Pb*Atg8 was not found to co-localize with these vacuoles [56]. In late erythrocytic stage parasites, heat stress led to cell death without typical signs of apopototic cell death, but with the appearance of cytoplasmic vacuoles suggestive of an autophagic type of cell death [77]. Prolonged amino acid starvation (4 h) in growing intracellular *T. gondii* led to mitochondrial fragmentation and loss of invasion competence; these two effects were blocked by the autophagy inhibitor, 3-MA [74]. Furthermore, both *Tg*Atg8 and *Pf*Atg8 exhibited a wider punctate distribution under these conditions [71,74]. Additionally, treatment with chloroquine in *P. falciparum* [78] or monensin in *T. gondii* [79] have both been proposed to lead to autophagy-related cell death. In particular, treatment with the anti-coccidial monensin led to *Tg*Atg8 puncta and mitochondrial degradation that was similarly blocked by 3-MA [79]. Taken together, all of these details are highly indicative of an autophagic cell death mechanism in response to certain cellular environments.

Much debate still exists over the role of the Atg proteins in Apicomplexa, particularly in *P. falciparum*. Surprisingly, one of the most contested roles for autophagy proteins is in canonical macroautophagy, as numerous studies have reported both the presence and absence of typical markers of autophagy [55]. Nevertheless, it has become increasingly clear that autophagy, even with a more limited repertoire of autophagy related proteins, has been adapted and repurposed to serve novel roles within the unique life cycles of Apicomplexa, as is the case with many fundamental processes in these parasites.

Finally, it is worth commenting on the interplay between parasite and host cell autophagy. From a variety of reports, it appears that host cell autophagy may serve both pro-life and pro-death functions for apicomplexan parasites. For example, recent results from the Pierce lab have shown that inhibiting mammalian target of rapamycin (mTOR) blocks the development of experimental cerebral malaria [80]. While mTOR is involved in several different conserved pathways regulating immune responses, which could be primarily responsible for this observation, it would nonetheless be interesting to determine the extent to which upregulation of host cell autophagy is involved in this effect.

Targeting the Plasmodial (and apicomplexan) Atg8-conjugation pathway

While much remains to be determined about the function of the Atg proteins within *Plasmodium* and the other Apicomplexa, one fact has become clear: whatever their functions, many of the Atg proteins are essential in the organisms that have been examined. As mentioned earlier, efforts to inhibit or knockdown Atg8 in particular apparently result in non-viable parasites [42,43,50]. Inhibiting the association of

Atg8 with parasite membranes, such as the apicoplast outer membrane, is predicted to be lethal to the cell, as Atg8 is required for proper cell function and Atg8's functional form within the parasite is its lipid-conjugated form. Therefore interfering with the conjugation pathway that leads to Atg8 lipidation, e.g. the Atg3-Atg8 interaction, is an attractive target for drug discovery. To obtain insights into the particular protein-protein interaction between *Pf*Atg3 and *Pf*Atg8, we co-crystallized *Pf*Atg8 with a peptide of *Pf*Atg3 [81]. The peptide (NDWLLPSY) was identified by sequence similarity to be the Atg8 family interacting motif of *Pf*Atg3, and it bound, as predicted, to the conserved W/L-site on *Pf*Atg8 [81]. This crystal structure (PDB ID: 4EOY), solved to 2.2 Å resolution, revealed the presence of a loop region in *Pf*Atg8 that does not exist in the homologous human or yeast Atg8 protein structures (Figure 1.2A) [81]. We termed this loop the apicomplexan loop, or A-loop, as it is conserved in apicomplexan Atg8s but not in human Atg8. It was further determined, through mutagenesis and deletion studies, that the A-loop is necessary for the interaction between *Pf*Atg3 and *Pf*Atg8. The existence and necessity of this loop makes *Pf*Atg8 a feasible target for drug discovery, as it provides the required divergence from homologous human structures to help prevent off-target effects. This point of divergence is important, since the main binding pocket on Atg8 (the W/L-site) is well conserved between *Plasmodium* and human Atg8 protein structures (Figure 1.2B). This structural knowledge allowed us to take two approaches towards novel lead discovery against the *Pf*Atg3-*Pf*Atg8 interaction.

First, the open access Medicines for Malaria Venture (MMV) Malaria box of 200 drug-like and 200 probe-like compounds was screened [82]. Using a Surface Plasmon Resonance (SPR) setup, we screened all 400 compounds against the *Pf*Atg3-*Pf*Atg8 interaction, identifying several compounds that met our criteria for sufficient inhibition of this protein-protein interaction [83]. Of these compounds, three shared a 4-pyridin-2-yl-1,3-thiazol-2-amine (PTA) scaffold and were found to inhibit the *Pf*Atg3-*Pf*Atg8 protein-protein interaction in a dose-dependent manner [83]. The most soluble of these hits was compound **1**, 2-methylsulfanyl-N-(4-pyridin-2-yl-1,3-thiazol-2-yl)benzamide, with an *in vitro* IC_{50} of 18.2 μM. Compound **1** bound directly to *Pf*Atg8 and, via virtual docking, was predicted to bind to the main binding pocket, the W/L-site (Figure 1.2C, D). The plasmodicidal activity of this compound was also tested against liver and blood stage parasites and showed an IC_{50} of 1.6 μM against blood stage *Plasmodium*. Treatment of infected liver cells for 96 h with 30 μM compound **1** led to a 50% decrease in the proportion of infected hepatocytes, while no effect was detected on liver cell viability. Of significant interest is the fact that treatment with compound **1** led to a dose-dependent upregulation of *Pf*Atg8 protein level along with a shift in lipidation state to primarily unlipidated *Pf*Atg8 as assayed via western blot. These data support the suggestion that compound **1** is acting against the predicted target and inhibits the *Pf*Atg8 and *Pf*Atg3 protein-protein interaction. Based on these promising results, we synthesized a 240-compound derivative library, using the PTA scaffold with a functional handle, in order to improve binding and specificity and to extend the molecule towards regions of the protein that are less conserved with the human homologues. In terms of the implications for drug development against other Apicomplexa, compound **1** showed growth inhibitory effects on *Cryptosporidium in vitro* [84], which suggests that it is also effective against the *Cp*Atg8-*Cp*Atg3 interaction. This result is not unexpected given the degree of conservation in the main binding pocket, which implies that the PTA compounds are also predicted to be active against other Atg3-Atg8 interactions in Apicomplexa.

In our second approach, we made direct use of the solved crystal structure to target the A-loop pocket of apicomplexan Atg8 using a virtual ligand screening approach [85]. Using the OpenEye suite of software (www.eyesopen.com), we virtually docked the ChemBridge Express library (www.chembridge.com) into the volume defined for the A-loop pocket (Figure 1.2C), specifying two protein contacts and two hydrogen bonds to anchor the hits within the pocket. From this initial screen, we cherry-picked 14 of the top hits based on visual inspection and cost. All of the hits produced some inhibition of the plasmodial protein-protein interaction and no inhibition of the homologous human

Atg8-Atg3 interaction when tested by SPR binding assays. apicomplexan Loop Compound 25 (ALC25) had the most potent effect and was used in follow-up assays to determine the *in vitro* IC_{50} against the *Pf*Atg3-*Pf*Atg8 interaction and to test against *P. falciparum* 3D7 blood and liver stage malaria. Its *in vitro* IC_{50} was 18.5 µM, and its effect on blood stage malaria was similar, with an IC_{50} of 19 µM. As was the case with PTA compound **1**, treatment of blood stage parasites with ALC25 led to an increase and shift in the lipidation state of *Pf*Atg8 as determined by western blot. ALC25 also showed activity against liver stage parasites, leading to a 30% decrease in the number of infected liver cells at 96 h post-treatment with 30 µM ALC25, while being well tolerated by human hepatocyte cell line. We also performed a secondary SPR assay to provide evidence for the predicted binding site of ALC25 within the A-loop pocket. With the established SPR assay [86], we immobilized a biotinylated version of the *Pf*Atg3 peptide to an SPR chip via NeutrAvidin capture – we used the same peptide that was co-crystallized with *Pf*Atg8 (PDB ID: 4EOY). The expected K_D between the peptide and *Pf*Atg8 was re-established and then it was demonstrated that ALC25 at inhibitory concentrations did not disrupt the peptide-*Pf*Atg8 interaction. As the peptide was shown to bind to the W/L-site in the crystal structure, this provides strong evidence that ALC25 binds to an alternative site. As suggested above, these A-loop compounds may have cross-reactivity against other Atg8 homologues in the same phylum since they target the conserved loop pocket. Either way, a similar approach would be likely to prove fruitful in generating novel leads against other apicomplexan Atg8 targets.

Both of our approaches resulted in the identification of novel compounds with activity *in vitro* and in parasite cultures against the *Pf*Atg3-*Pf*Atg8 interaction. Our next step is to optimize these leads to provide increased binding affinity and specificity. Interestingly, there is the possibility of linking these two molecules, or these two sites, as the functional handle of the PTA compound derivatives is predicted to point toward the A-loop pocket (Figure 1.2D). A drug that would bind to both of these sites could have the exact desired qualities of increased specificity and affinity. As mentioned above, the greater degree of conservation within apicomplexan Atg8 homologues supports our hypothesis that the methods suggested here may have broader implications for the discovery of new intervention strategies against Apicomplexa in general.

Concluding Remarks

The last few years have seen a substantial increase in what is known about apicomplexan autophagy and the parasites' interactions with their host cells. As apicomplexan autophagy represents a conserved and essential pathway in these parasites, it is of significant importance on two fronts. First, the sheer fact of its necessity to parasite function indicates that autophagy proteins within these organisms have been repurposed to serve novel functions, as many autophagy proteins are not essential within other higher-level eukaryotes [13,26]. This hypothesis is well supported by the highly documented association of Atg8 with the apicoplast, as an association with non-transient organelles is well outside of the extent of classical macroautophagy. Second, as was made clear in the third section of this review, this conserved pathway, and the Atg3-Atg8 interaction in particular, represent fertile ground for the discovery of novel anti-apicomplexan small molecule interventions. Two such small molecules inhibiting the *Pf*Atg3-*Pf*Atg8 interaction were described above with evidence to support the theory that these molecules, and the methods by which they were discovered, could have broader applications within the Apicomplexa phylum as a whole [81,83] (Hain 2016).

As more is learned about this fundamental pathway, it is likely that new targets and intervention strategies will arise, with an emphasis on seeking low-cost solutions to the global health issues presented by Apicomplexa. Such low-cost interventions are of critical importance on the global scale, as many of the countries continually plagued by malaria and other apicomplexan diseases do not have the financial means to afford the high cost therapies that arise from pharmaceutical brute force drug discovery. For this reason, virtual docking will continue to be a useful tool within this field, as making use of computational tools could be expected to reduce drug development costs by up to 50 percent [87].

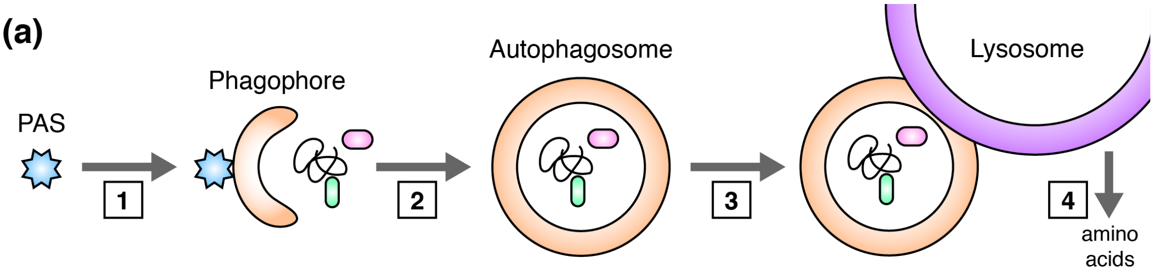
On a final note, it is also worthwhile to comment on the connection between this research and potential applications to human autophagy interventions. Defects in autophagy have been linked to a number of important human diseases, including immune disorders [88], heart disease [89], neurodegenerative diseases [90], and most notably, cancer [91,92]. As appears to be a repeating trend, autophagy has been found to serve both pro-survival and pro-death functions within various cancer types [91,92]. Due to its pro-survival role, autophagy has recently become an important target for new cancer therapies. Because of this, our research has had potential applications to human cancer, a connection also made for other antimalarials [93,94]. For example, in the process of screening the VLS compounds against the human LC3-Atg3 interaction, we discovered a compound that would not be appropriate as an antimalarial due to its activity against the human interaction [85]. As regions of Atg8 are well conserved between *Plasmodium* and human homologues, it is only natural that small molecule screening against one target will have the added benefit of potentially uncovering inhibitors against the other target.

Figures

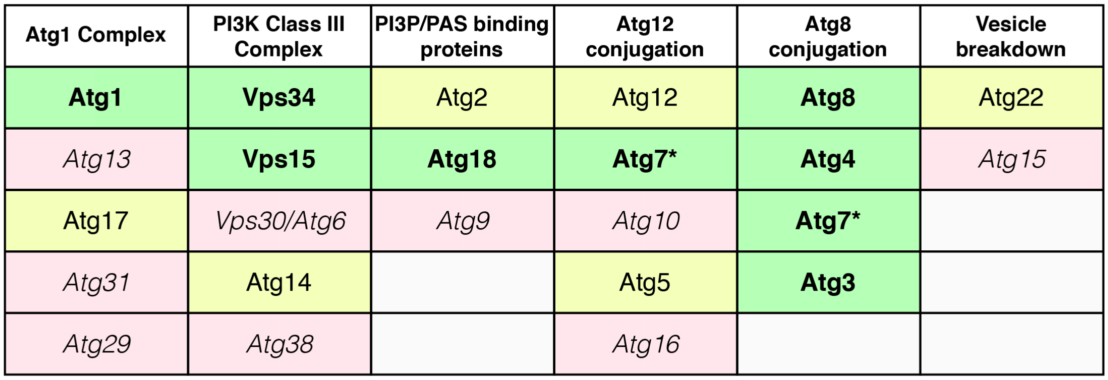
Figure 1.1. An overview of canonical macroautophagy in yeast and *Plasmodium*. (A) Step 1: induction and phagophore assembly. The Atg1 complex forms at the PAS, leading to recruitment and activation of the PI3K class III complex. The PI3K complex converts PI to PI3P in the phagophore membrane resulting in Atg2, Atg18, and Atg9 binding. Step 2: phagophore elongation and closure to form the autophagosome. This step relies on Atg8 being conjugated to PE. Atg8–PE is the product of the Atg8 conjugation pathway, the last step of which is facilitated by Atg12–Atg5-Atg16. Atg12–Atg5 is the product of the Atg12 conjugation pathway. Step 3: autophagosome fuses with the lysosome to form an autolysosome. Step 4: digestion of autolysosomal contents and release of recycled amino acids. Atg15 and Atg22 facilitate breakdown of the membrane-bound vesicle within the autolysosome to expose the contents for digestion. The products of digestion, amino acids, are then released into the cell. (B) Conservation of canonical macroautophagy proteins within *Plasmodium* [17,55]. Bold text on a green background indicates a conserved homologue exists; regular text on a yellow background, potential homologues identified; italics on a red background, no detectable homologue. * – Atg7 is involved in both Atg8 and Atg12 conjugation pathways.

Figure 1.1

(a)



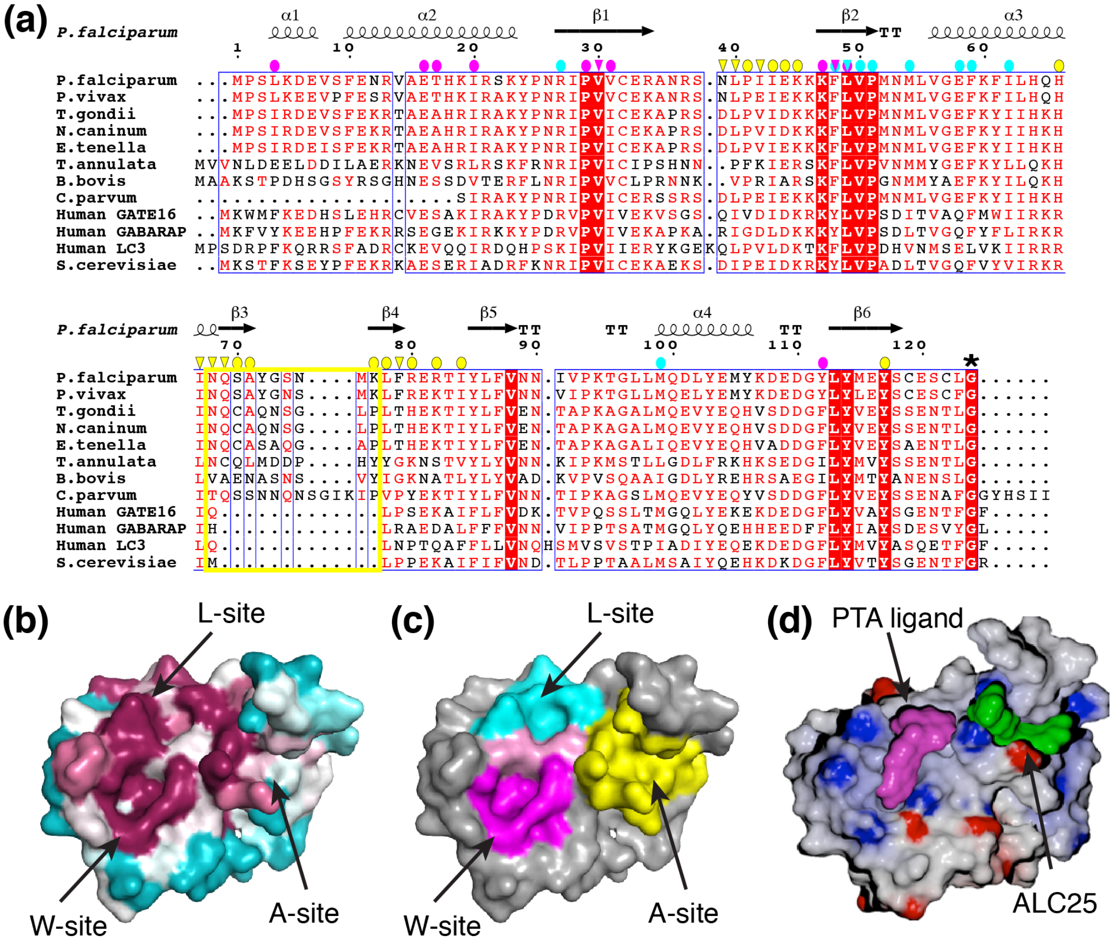
(b)



Atg1 Complex	**PI3K Class III Complex**	**PI3P/PAS binding proteins**	**Atg12 conjugation**	**Atg8 conjugation**	**Vesicle breakdown**
Atg1	**Vps34**	Atg2	Atg12	**Atg8**	Atg22
Atg13	**Vps15**	**Atg18**	**Atg7***	**Atg4**	*Atg15*
Atg17	*Vps30/Atg6*	*Atg9*	*Atg10*	**Atg7***	
Atg31	Atg14		Atg5	**Atg3**	
Atg29	*Atg38*		*Atg16*		

Figure 1.2. *Plasmodium* and apicomplexan Atg8 proteins. (A) Multiple sequence alignment of various apicomplexan, human, and yeast Atg8 homologues generated by ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Visualized with ESPript 3.0, with global score set to 0.4 (espript.ibcp.fr). Ovals indicate side chain contacts in the three pockets on *Pf*Atg8; triangles indicated backbone contacts in these pockets. Magenta corresponds to the W-site pocket; cyan, the L-site pocket; and yellow, the A-loop pocket. The A-loop itself is outlined in yellow, revealing conservation of the A-loop within apicomplexan, but not in human or yeast homologues. The C-terminal glycine is indicated by a black asterisk, illustrating that this glycine is exposed in most apicomplexan Atg8 homologues, but not in *C. parvum*, human or yeast Atg8. (B) Conservation of pockets on *Pf*Atg8 (PDB ID 4EOY) generated using Consurf with the default values (consurf.tau.ac.il). The W- and L-sites are highly conserved, while the A-loop pocket is not. (C) Structure of *Pf*Atg8 color-coded as in (A): magenta, W-site; cyan, L-site; yellow, A-loop pocket. The pink region between the W- and L-sites indicates residues involved in both sites. (D) Electrostatic surface representation of *Pf*Atg8 rendered using OpenEye Vida and POV-Ray (povray.org). Predicted binding modes of one of the PTA ligands, bound in the W- and L-sites, and ALC25, bound in the A-loop pocket. From these poses, it appears there is the potential to link the two sites.

Figure 1.2


(a)
P.falciparum
α1 α2 β1 β2 TT α3
1 10 20 30 40 50 60
P.falciparum ...MPSLKDEVSFENRVAETHKIRSKYPNRIPVVCERANRS.NLPIIEKKKFLVPMNMLVGEFKFILHQH
P.vivax ...MPSLKEEVPFESRVAETHKIRAKYPNRIPVVCEKANRS.NLPEIEKKKFLVPMNMLVGEFKFILHQH
T.gondii ...MPSIRDEVSFEKRTAEAHRIRAKYPNRIPVICEKAPRS.DLPVIDKKKFLVPMNMLVGEFKYIIHKH
N.caninum ...MPSIRDEVSFEKRTAEAHRIRAKYPNRIPVICEKAPRS.DLPVIDKKKFLVPMNMLVGEFKYIIHKH
E.tenella ...MPSIRDEISFEKRTAEAHRIRAKYPNRIPVICEKAPRS.DLPVIEKKKFLVPMNMLVGEFKYIIHKH
T.annulata MVVNLDEELDDILAERKNEVSRLRSKFRNRIPVICIPSHNN..PFKIERSKFLVPVNMMYGEFKYLLQKH
B.bovis MAAKSTPDHSGSYRSGHNESSDVTERFLNRIPVVCLPRNNK..VPRIARSKFLVPGNMMYAEFKYILQKH
C.parvum ...................SIRAKYPNRIPVICERSSRS.DLPEIEKKKFLVPMNMLVGEFKYIIHKH
Human GATE16 ..MKWMFKEDHSLEHRCVESAKIRAKYPDRVPVIVEKVSGS.QIVDIDKRKYLVPSDITVAQFMWIIRKR
Human GABARAP ..MKFVYKEEHPFEKRRSEGEKIRKKYPDRVPVIVEKAPKA.RIGDLDKKKYLVPSDLTVGQFYFLIRKR
Human LC3 MPSDRPFKQRRSFADRCKEVQQIRDQHPSKIPVIIERYKGEKQLPVLDKTKFLVPDHVNMSELVKIIRRR
S.cerevisiae ..MKSTFKSEYPFEKRKAESERIADRFKNRIPVICEKAEKS.DIPEIDKRKYLVPADLTVGQFVYVIRKR
P.falciparum
β3 β4 β5 TT TT α4 TT β6
70 80 90 100 110 120
*
P.falciparum INQSAYGSN....MKLFRERTIYLFVNN.IVPKTGLLMQDLYEMYKDEDGYLYMEYSCESCLG......
P.vivax INQSAYGNS....MKLFREKTIYLFVNN.VIPKTGLLMQELYEMYKDEDGYLYLEYSCESCFG......
T.gondii INQCAQNSG....LPLTHEKTIYLFVEN.TAPKAGALMQEVYEQHVSDDGFLYVEYSSENTLG......
N.caninum INQCAQNSG....LPLTHEKTIYLFVEN.TAPKAGALMQEVYEQHVSDDGFLYVEYSSENTLG......
E.tenella INQCASAQG....APLTHEKTIYLFVEN.TAPKAGALIQEVYEQHVADDGFLYVEYSAENTLG......
T.annulata LNCQLMDDP....HYYGKNSTVYLYVNN.KIPKMSTLLGDLFRKHKSEDGILYMVYSSENTLG......
B.bovis LVAENASNS....VYIGKNATLYLYVAD.KVPVSQAAIGDLYREHRSAEGILYMTYANENSLG......
C.parvum ITQSSNNQNSGIKIPVPYEKTIYLFVNN.TIPKAGSLMQEVYEQYVSDDGFLYVEYSSENAFGGYHSII
Human GATE16 IQ............LPSEKAIFLFVDK.TVPQSSLTMGQLYEKEKDEDGFLYVAYSGENTFGF.....
Human GABARAP IH............LRAEDALFFFVNN.VIPPTSATMGQLYQEHHEEDFFLYIAYSDESVYGL.....
Human LC3 LQ............LNPTQAFFLLVNQHSMVSVSTPIADIYEQEKDEDGFLYMVYASQETFGF.....
S.cerevisiae IM............LPPEKAIFIFVND.TLPPTAALMSAIYQEHKDKDGFLYVTYSGENTFGR.....
(b)
L-site
W-site
A-site
(c)
L-site
W-site
A-site
(d)
PTA ligand
ALC25

CHAPTER 2

Virtual Screening and Experimental Validation Identify Novel Inhibitors of the *Plasmodium falciparum* Atg8–Atg3 Protein-Protein Interaction

Sections previously published in:

Miller, A.S.*, Hain, A.U.P.*, Levitskaya, J., and Bosch, J. (2016) Virtual Screening and Experimental Validation Identify Novel Inhibitors of the *Plasmodium falciparum* Atg8 – Atg3 Protein – Protein Interaction. *ChemMedChem*. (* – authors contributed equally to this work.)

Acknowledgements

I acknowledge Dr. Hain for her blood-stage parasite work and Dr. Levitskaya for her liver-stage parasite work. I also acknowledge Dr. Bosch and Dr. Hain for the initial compound selection. In the following work, I was responsible for all *in vitro* assays (SPR, TSA, ITC) and for the sequence/structural comparisons (MSA, etc.). This work is reproduced with permission under the ACS AuthorChoice License, which permits copying/redistribution of the article for non-commercial purposes.

Published acknowledgements: We thank the members of the Bosch research group for helpful discussions and comments on the manuscript. We acknowledge the use of the Johns Hopkins Malaria Insectary and Parasitology Core Facility and the Genetics Resources Core Facility. Part of the research is funded by a pilot grant from the Johns Hopkins Malaria Research Institute (JHMRI) (J.B.), a pre-doctoral fellowship from the JHMRI (A.U.P.H.), and a training grant T32M008403-24 (A.S.M.).

Abstract

New therapies are needed against malaria, a parasitic infection caused by *Plasmodium falciparum*, as drug resistance emerges against the current treatment, artemisinin. We previously characterized the Atg8–Atg3 protein-protein interaction (PPI), which is essential for autophagy and parasite survival. Herein we illustrate the use of virtual library screening to selectively block the PPI in the parasite without inhibiting the homologous interaction in humans by targeting the A-loop of *Pf*Atg8. This A-loop is important for Atg3 binding in *Plasmodium*, but is absent from the human Atg8 homologues. In this proof-of-concept study, we demonstrate a shift in lipidation state of *Pf*Atg8 and inhibition of *P. falciparum* growth in both blood- and liver-stage cultures upon drug treatment. Our results illustrate how *in silico* screening and structure-aided drug design against a PPI can be used to identify new hits for drug development. Additionally, as we targeted a region of Atg8 that is conserved within apicomplexans, we predict that our small molecule will have cross-reactivity against other disease-causing apicomplexans, such as *Toxoplasma*, *Cryptosporidium*, *Theileria*, *Neospora*, *Eimeria*, and *Babesia*.

Introduction

Malaria continues to be a major public health and economic burden in the developing world, resulting in over half a million deaths each year, most of which are caused by the *Plasmodium falciparum* species of the malaria parasite [95]. Currently, artemisinin combination therapy is the first-line defense for *P. falciparum* malaria in most countries. As was the case for previous antimalarials such as chloroquine (CQ), resistance is emerging against artemisinin and has been detected in five countries in Southeast Asia [36,95]. It is critical to have new and effective therapies ready as resistance inevitably spreads to other parts of the world.

Plasmodium is a member of the phylum Apicomplexa, a group of parasitic protists characterized by their possession of an apical complex involved in host cell invasion. These organisms also frequently contain a unique plastid organelle called the apicoplast. A number of apicomplexan organisms cause parasitic infections that pose a threat to young children and immune-suppressed individuals. For example, *Toxoplasma gondii*, the causative agent of toxoplasmosis, asymptomatically infects up to a third of the world's population and, though harmless in healthy individuals, can be fatal when the immune system becomes compromised [96]. Additionally, species of *Cryptosporidium* are responsible for the majority of parasitic protozoan water-borne outbreaks globally [32], with limited treatments for young, malnourished, or HIV-infected patients [97].

Apicomplexans pose a threat not only to human health. *Babesia*, *Eimeria*, *Neospora*, and *Theileria* species are important veterinary pathogens in countries across the globe, resulting in significant economic burdens to the livestock and poultry industries. *Neospora caninum*, for example, is a significant contributor to cattle abortion, resulting in losses of over $500 million annually in the United States dairy industry [35].

Recent work has shown the autophagy pathway is conserved and essential in both *P. falciparum* and *T. gondii* [47,48,50,98]. While in higher-level eukaryotes, autophagy is primarily a degradation pathway for bulk cytoplasm, protein aggregates, and damaged organelles, apicomplexans have adapted this pathway to serve diverse purposes specific to the parasite lifecycle [43,45,47,48,50,54,98]. Because the autophagy machinery in these organisms is more conserved across apicomplexan organisms than between apicomplexans and humans, we identified a target that would allow for the inhibition of autophagy across parasite species. Our first proof-of-concept study was recently published with data on *P. falciparum* [81,83]. Our previous hit, *N*-(4-methylphenyl)-4-pyridin-2-yl-1,3-thiazol-2-amine, identified from the Medicines for Malaria Venture Malaria Box [82], was also shown to have growth inhibitory effects on *Cryptosporidium* cultures [84], suggesting that the *Cp*Atg8–*Cp*Atg3 interaction might be required in this organism, as well.

Atg8 is an essential component of the core autophagy machinery. In yeast and other eukaryotic organisms, its main role is in the expansion and closure of the autophagic vesicle, or autophagosome, which sequesters material to be degraded in the vacuole [17,46]. Atg8 is covalently attached to the autophagosome membrane through a series of intermediate thioester formations with its E1-activating enzyme, Atg7, and E2-conjugating enzyme, Atg3. The biologically relevant form of Atg8 is thus its lipidated form, which has been shown to be the predominant species within *P. falciparum* parasites; however, evidence suggests that *Pf*Atg8 does not play the same role within autophagy, despite its essentiality [44,45,53,54].

In order to lipidate Atg8, Atg3 interacts with two conserved hydrophobic pockets on Atg8, termed the W- and L-sites. We previously identified an additional region in *Plasmodium* Atg8 that is important for the interaction with plasmodial Atg3. This region was termed the A-loop, as it is conserved in apicomplexans and absent in human homologues (Figures 2.1 and 2.2) [81]. This divergence provides the opportunity for specificity and selectivity in developing anti-apicomplexan inhibitors.

Virtual ligand screening (VLS), an *in silico* method in which compounds are computationally docked to a known protein structure, is increasingly used in the drug development field [99]. We used the X-ray structure of *P. falciparum* Atg8, PDB ID: 4EOY [81], as a template for the VLS. Using this method, we identified compounds that inhibit the Atg8–

Atg3 interaction in *P. falciparum* without inhibiting the homologous interaction in humans. Because of the sequence conservation to other disease-causing parasites, such as *Toxoplasma* and *Cryptosporidium*, we predict that inhibitors developed against this target could be cross-reactive with other parasites. Targeting protein regions conserved across disease-causing organisms using structure-aided drug design has the potential to identify compounds with cross-species reactivity.

Results and Discussion

Virtual docking used to select 14 compounds for testing

To identify potential *Pf*Atg8–*Pf*Atg3 inhibitors, we undertook VLS against the co-crystal structure of *Pf*Atg8 with a peptide of *Pf*Atg3 (PDB ID: 4EOY). We removed the peptide and used the structure of *Pf*Atg8 as the receptor in docking. Multiple sequence alignment and structural analysis revealed a region of Atg8, termed the A-loop, which is conserved within Apicomplexa, but absent from human homologues (Figures 2.1 and 2.2). Therefore, docking was constrained to the A-loop pocket (Figure 2.3A,B). To constrain it to the A-loop, we specified that the small molecules have two backbone hydrogen bond interactions within the A-loop (His67, Ile68), and two side chain residues (Lys46 and Glu45) were selected as contacting residues for the docking algorithm to ensure that the search was targeted to the specific pocket of interest. The His67 and Ile68 backbone interactions are sequence independent; their selection was intended to bias the search toward hits that could have an effect on multiple apicomplexan Atg8 homologues. These residues have the added benefit of being absent from the human homologues, which increases the chances for limited off-target effects.

The ChemBridge library (July 2012) contained 369,632 descriptors for small molecules, from which 10,848 were considered toxic or reactive using the default filter algorithm of the program FILTER from the OpenEye suite [100]. A property characterization of the filtered VLS library is represented in Supporting Information Figure 2.S1A-C. The final conformer library used for docking studies contained on average 816 conformers per small molecule, resulting in a total of 546,161,945 docking trials tested by FRED (Supporting Information Figure S2.1D) [101]. The default Chemgauss 4 scoring algorithm was used to rank the hits from VLS. The hits were further ranked according to known toxicity data, predicted solubility, balance between polar/apolar interactions, and to achieve maximal diversification. Of 29 molecules selected based on this ranking, we were able to obtain 14 from commercial sources for further *in vitro* and *in vivo* parasite testing (Table 2.1). Figure 2.3C summarizes the workflow for this selection process.

Surface Plasmon Resonance (SPR) interaction assay confirms in vitro inhibition by apicomplexan loop compound 25 (ALC25)

We tested the 14 VLS hits for *in vitro* inhibition of the plasmodial Atg8–Atg3 interaction using our established SPR competition assay (Figure 2.4A) [81]. Binding is measured by SPR response of an injection of purified *Pf*Atg8 alone or in the presence of small molecules over immobilized *Pf*Atg3; 13 of the 14 compounds inhibited the *Pf*Atg3-*Pf*Atg8 interaction (Figure 2.4B).

The molecules were further tested for inhibition of the homologous human interaction between hLC3 and hAtg3. None of the compounds inhibited hAtg3-hLC3 binding (Figure 2.4C). We observed that some of the compounds (ALC11, ALC13, ALC19, ALC20, and ALC30) either precipitated or caused protein precipitation at the concentration tested in the assay. These compounds are indicated in blue in Figure 2.4B,C and were not further pursued. The effect of the molecules on the plasmodial and human interactions was plotted in a two-species inhibitor/stabilizer (TSIS) plot for ease of comparison (Figure 2.4D).

Dose dependency, direct binding, and interaction site of ALC25 confirmed in vitro by SPR, thermal stability assay, and ITC

ALC25 and ALC24, the best hits, were chosen for verification of dose-dependent inhibition of the interaction. ALC24 led to a notable increase in binding for the human interaction, suggesting that it would not be viable as an antimalarial. However, because ALC24 interferes with the human interaction by significantly stabilizing the binding of hLC3 to hAtg3, it may have applications in certain cancer therapies, as autophagy has been linked to cancer survival, proliferation, and growth [91,92]. In terms of the plasmodial Atg8–Atg3 interaction, of the two compounds, only ALC25 led to dose-dependent inhibition of plasmodial Atg8–Atg3, with an SPR IC_{50} value of 18.5±2.1 μM (Figure 2.5A).

Direct binding of ALC25 was tested using thermal shift assays on the protein *Pf*Atg8^{CM}, a construct with three cysteine residues mutated to isoleucine or serine for stability (the construct for the crystal structure [81]). ALC25 led to a dose-dependent decrease in the T_m, which suggests direct binding to *Pf*Atg8, but also indicates that the compound may not be well-suited for co-crystal matrix screening (Figure 2.5B). This decrease in T_m could be the result of engaging Lys46 (one of the docking constraints) away from the W-site, leading to a slight destabilization of the protein core.

To confirm that ALC25 does not bind to the W/L site of *Pf*Atg8, we used an established SPR assay [86] to determine whether ALC25 interferes with the binding of the *Pf*Atg3 peptide to *Pf*Atg8. The *Pf*Atg3 Atg8-interacting-motif peptide (NDWLLPSY) was previously co-crystallized with *Pf*Atg8 and shown to bind in the W/L site of *Pf*Atg8 (PDB ID: 4EOY) [81]. Avitag-*Pf*Atg3 peptide was biotinylated and immobilized on an SPR chip coated with NeutrAvidin. To validate that our Avitag-*Pf*Atg3 chip was functional, we passed recombinant *Pf*Atg8 over the chip and observed binding and dissociation from the chip, which were in agreement with our previously reported results [81]. To assess whether or not ALC25 binds to the W/L site, we passed *Pf*Atg8 over the *Pf*Atg3 peptide in the presence of 1 mM ALC25 and observed no inhibition of the interaction (Figure 2.5C). This, along with our docking results, suggests that ALC25 binds predominantly to the apicomplexan loop region on *Pf*Atg8. Of note, ALC25 had the second-highest ranking overall in the original docking screen. The docking pose of ALC25 and its interactions with the pocket are shown in Figure 2.6.

While this SPR assay confirmed that ALC25 does not bind to the W/L site, we used ITC to show that the loop itself is involved in binding. We titrated 50 μM ALC25 into a solution of 5 μM *Pf*Atg8^{CM} and *Pf*Atg8$^{CM/LD}$, which is identical to the CM construct except for a deletion of nine amino acids involved in the A-loop: residues 69-77 (NQSAYGSNM) [81]. Although *Pf*Atg8 is not well-suited to ITC due to its low solubility, we nevertheless observed heat due to binding that could be fit to a one-site binding model for the wild-type protein, whereas we observed no appreciable binding for the loop deletion mutant, thus confirming the role of the A-loop in ALC25 binding (Supporting Information Figure S2.2).

ALC25 exhibits modest inhibition in P. falciparum cultures

We measured growth inhibition of *P. falciparum* 3D7 blood-stage cultures by ALC25 using the SYBR green I assay, which measures binding of a fluorescent dye to parasite DNA [102]. In our assay, the well-established antimalarial CQ had an IC_{50} value of 3±2 nm. ALC25 was less potent, with an IC_{50} value of 19±6 μM (Figure 2.7A).

In other species, drug inhibition of autophagy leads to an increase in Atg8 protein levels, as Atg8 cannot be degraded and the cell up-regulates Atg8 expression [103]. To confirm that ALC25 was affecting *Pf*Atg8 in the parasite, we probed the lipidation status and protein level of *Pf*Atg8 with a 4-20^% acrylamide gel and immunoblot analysis. After a five-hour treatment with 100 μM ALC25, *Pf*Atg8 protein levels increased over 10-fold relative to the DMSO-treated control. Additionally, the mobility of Atg8 was retarded, indicative of the unlipidated Atg8 species (Figure 2.7B,C) [104,105].

It has previously been reported that *Pf*Atg8 partially localizes to the apicoplast during exoerythrocytic development of *Plasmodium* in hepatocytes [56]. As the apicoplast is an essential organelle for *Plasmodium* survival, we hypothesized that Atg8 could serve an important role in growth during the liver stage. We next assessed the effect of ALC25 on the *in vitro* development of *P. falciparum* exoerythrocytic forms (EEFs) using a transgenic parasite strain expressing GFP throughout its life cycle [106]. While ALC25 had little effect on GFP fluorescence intensity, the number of infected HC-04 cells detected by flow cytometry 96 hours post-infection was decreased by 30^% in hepatocyte cultures treated with 30 μM ALC25 (Figure 2.8A). In contrast, ALC25 did not affect cell viability of HC-04 cells after a 96-hour treatment (Figure 2.8B).

Conclusions

Autophagy is a catabolic process performed by eukaryotic cells in order to maintain homeostasis and to degrade unwanted or toxic cellular content, such as misfolded proteins [107]. The autophagy pathway has been demonstrated to be important for various functions in yeast and humans [107,108], but studies in the past five years have also shown that autophagy is necessary in human pathogens, such as *Plasmodium* [17,44,53].

In the process of drug discovery, promiscuity, pleiotropy, and polypharmacology continue to be a problem. Many of the drugs currently used as autophagy inhibitors, such as 3-methyladenine or CQ, are promiscuous compounds. Other autophagy modulators, such as rapamycin and wortmannin, are inherently pleiotropic as their protein targets (mTor and VPS34/Beclin1, respectively) are effectors of multiple pathways [109–111]. Beclin1, in particular, may have an autophagy-independent function [112,113]. Furthermore, polypharmacology, the interaction of drugs with multiple protein targets, is a concern for many FDA-approved drugs [114].

Related to these concerns, current intervention methods targeting the autophagy pathway are somewhat indirect and intervene very early in the process of inducing autophagy, which may allow the cell to regulate the autophagy pathway through later checkpoint events. Targeting downstream events is attractive in circumventing this potential pitfall. Autophagy inhibitors that specifically prevent autophagosome expansion have been reported in humans and are being considered for anticancer drug therapy [115].

Our recently identified antimalarial compounds with a pyridinylthiazolamine (PTA) scaffold inhibit the critical phosphatidylethanolamine (PE)-conjugation step of LC3/Atg8 by its E2-conjugating enzyme, Atg3.[83] Our PTA compounds were designed to interfere with the Atg8–Atg3 protein-protein interaction (PPI) and thereby prevent the lipidation of Atg8. Only this lipidated form of Atg8 is active: it participates in elongation of the autophagosomal membrane in higher eukaryotes, and is the predominant species in *Plasmodium* [44,45,53,54].

Our bioinformatics analysis suggests that the autophagy pathway is conserved in multiple human and livestock pathogens and likely fulfills important functions in these organisms.

We used an *in silico* approach to target the A-loop pocket of *Pf*Atg8. This approach led to the discovery of a compound that inhibits the *in* *vitro* Atg8–Atg3 interaction in *Plasmodium* and shows activity in blood and liver stages. Although the potency is in the micromolar range, ALC25 was well tolerated by a human hepatocyte cell line at these concentrations. The IC_{50} values of ALC25 against the *in vitro* PPI and against blood-stage parasite growth *in vitro* are in good agreement, suggesting that ALC25 may target the Atg8–Atg3 interaction *in vivo*. In support of this, ALC25 led to accumulation of the unlipidated Atg8 species as quantified by western blot with specific antibodies, a phenomenon previously observed for another inhibitor of this PPI [83]. We now have the opportunity to proceed with lead optimization of a small molecule derived from VLS based on *in vitro* validation and apparent on-target effect in parasite cultures. In addition, we show that two independent drug discovery pathways result in different small-molecule entities capable of interfering with the process of Atg8 lipidation.

It remains to be tested whether ALC25 has cross-reactivity against other apicomplexan species. However, our bioinformatics analysis suggests ALC25 may have activity against *Eimeria*, *Toxoplasma*/*Neospora*, *Cryptosporidum*, *Theileria*, and *Babesia* (Figure 2.2C).

The importance and feasibility of targeting PPI interfaces has become increasingly clear over time [116,117]. In this case, we demonstrate that the development of PPI inhibitors is a promising and viable strategy for future lead identification using orthogonal screening and validation methods. Our structural models support the hypothesis that it may be possible to bridge the W/L site, occupied by the PTA molecule, with the A-site, occupied by ALC25. A molecule that interacts with both sites would be predicted to have increased potency and provide selectivity toward the parasitic Atg8 proteins, as the A-loop is absent in human homologues.

Materials and Methods

Bacterial strains and culture conditions. *Escherichia coli* strains were cultured in Terrific Broth (TB) medium supplemented with 50 µg/mL ampicillin or kanamycin as appropriate. All proteins were expressed using the *E. coli* Rosetta 2 strain, which has an additional plasmid for expressing rare codons with a chloramphenicol-resistance cassette; cultures were thus additionally supplemented with 34 µg/mL chloramphenicol.

Plasmid and strain construction. The plasmids and strains used to express hLC3, *Pf*Atg8, and *Pf*Atg3 have been described previously [81]. hAtg3 was received from the Johns Hopkins Genetic Resources Core Facility on a pENTR221 vector. The gene was PCR amplified using custom primers 5'-CATG CCAT GGGC CAGA ATGT GATT AATA CTGT GAAG GGAA (sense) and 5'-cgcg gatc cgtt acat tgtg aagt gtct tgtgt agtc (antisense). The PCR fragment was digested with New England BioLabs® (NEB) *Nco*I and BamHI restriction enzymes, and then ligated with NEB

T4 DNA ligase into a modified pRSF vector containing an N-terminal maltose binding protein (MBP) tag followed by a tobacco etch virus (TEV) protease cleavage site [86]. The plasmid was then transformed into chemically competent *E. coli* Rosetta 2 cells.

Protein expression and purification. For hLC3, *Pf*Atg8, and *Pf*Atg3, expression and purification were performed as previously described [81]. After transformation of hAtg3, a single colony was selected and sequence verified. Cells from this stock were grown in TB media with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. Cells were grown in the presence of 1% glucose at 37°C, shaking at 250 rpm until an OD_{600} of 3.0 was reached. At this point, the cells were induced with 0.5 mM IPTG and then continued growing overnight at 20°C with shaking. All steps after overnight incubation were conducted at 4°C. The cells were harvested by centrifugation at 1300 *g* for 45 min and resuspended in lysis buffer: 10 mM HEPES (pH 7.5), 150 mM NaCl, 1 mM DTT with one tablet of Roche® Complete EDTA-free protease inhibitor per 100 mL and 10 U benzonase (Sigma-Aldrich®). Using an Avestin® EmulsiFlex-C5 cell disruptor, the cells were lysed at 15 kpsi (100 MPa). This lysed solution was cleared at 18,000 *g* for 45 min, after which the supernatant was decanted to incubate with amylose resin (NEB) for 30 min. The resin was washed with lysis buffer before the protein was incubated for 30 min with elution buffer (lysis buffer with 10 mM maltose). Eluted hAtg3 was cleaved with homemade TEV protease at a mass ratio of 1:100, while simultaneously dialyzed into 10 mM HEPES (pH 7.5), 100 mM NaCl, 1 mM DTT. After cleaving, hAtg3 was separated from MBP and TEV protease by anion exchange using a RESOURCE™ Q column (GE® Healthcare) on an ÄKTA™ purifier system. Purified proteins were analyzed by SDS-PAGE and quantified using a Nanodrop (ThermoFisher®) spectrophotometer at λ=280 nM using the molar extinction coefficient for each of the proteins as calculated from the primary sequence by ProtParam [118]. For SPR, proteins were concentrated to 1-3 mg/mL and buffer exchanged into the indicated running buffer.

VLS with FRED. The OpenEye suite of programs [100] was used to prepare an *in silico* library of the ChemBridge commercial small-molecule database. Structure description files (.sdf) were obtained from ChemBridge and first subjected to the OpenEye FILTER program for removal of small molecules with known toxic and reactive groups. Next, a conformer library was generated using OMEGA2 [119] with a maximum of 2000 conformers per small molecule and otherwise using default parameters. This conformer library was then docked against the co-crystal structure of *Pf*Atg8–*Pf*Atg3, PDB ID: 4EOY [81], after removing the *Pf*Atg3 peptide and water molecules. The protein receptor was generated with OpenEye's make_receptor program using default values and adding restraints in the form of two backbone interactions and two side chain contacts within the apicomplexan loop pocket to serve as anchoring points for the docking (Figure 2.3A,B). Docking was carried out on a 24 core MacPro using FRED [101] with default settings using a fine grid of 0.5 Å. Visualization and inspection of the top-500 docking results was carried out with OpenEye's visualization software, VIDA. A selection of 29 molecules were cherry picked, from which 14 were readily available from commercial sources for testing in biochemical and biophysical assays (Table 2.1).

SPR PPI assay. All measurements were carried out on a BiaCore 3000 instrument at 25°C. We used the previously described SPR competition assay [81] to test inhibition of the *Pf*Atg8–*Pf*Atg3 and hLC3-hAtg3 interactions. The standard running buffer was 10 mM HEPES, 150 mM NaCl, 0.01% Tween 20, adjusted to pH 7.5. Prior to performing SPR assays, all solutions were sterile filtered and degassed for 1 h at room temperature. For small-molecule inhibitor studies, the same running buffer was supplemented with 1% DMSO. All dose-dependency measurements were carried out in triplicate and corrected for DMSO absorption effects using double referencing with interspersed blank injections and an untreated flow cell on the SPR chip and the SPR analysis software Scrubber 2.0 (BioLogic™). For each interaction pair, a new CM5 SPR chip was used. To exclude surface decay during the experiment, positive control injections were interspersed every 12th injection. For all interaction studies, the Atg3 homologue was immobilized, and the corresponding Atg8 partner was passed as an analyte over the chip in either the presence or absence of small-molecule inhibitors, dissolved to 100 mM in 100% DMSO and tested at 500 µM final concentration. Prior to the immobilization of Atg3 to an SPR chip, a pH scouting from pH 6.0 to pH 4.0 in intervals of 0.5 pH units was carried out to determine the best pre-concentration conditions for immobilization. Briefly, flow cells of an SPR chip were activated using *N*-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) for 7 min at a flow rate of 20 µL/min. For *Pf*Atg3 immobilization, a solution of 50 µg/mL *Pf*Atg3 in 10 mM sodium acetate (pH 5.0) was passed over one flow cell until >500 response units (RU) were captured. The same procedure was followed for the immobilization of hAtg3. Activated free binding sites on the chip were blocked using 1M ethanolamine for an additional 7 min. All experiments were carried out at 25°C at a 40 µL/min flow rate and 75 s contact time, followed by a 1 min dissociation time and a 25 µL regeneration injection of 2M $MgCl_2$. The dose-response curves were used to determine the approximate concentration of analyte (*Pf*Atg8 or hLC3) needed to obtain a particular RU value, typically in the range of 100-150 RU. A fixed amount of analyte was then passed over the chip resulting in at least 100-150 RU in the absence of small molecules.

TSIS plot. To determine which inhibitors to pursue after the initial screen, it is useful not only to consider the variance of the control response (to determine if the inhibited response differs significantly from the control), but also to consider the inherent variability in the non-control responses, which is a function of many features, including the density of active protein on the chip. To do this, we considered the standard deviation of the non-control responses as a means by which to detect compounds that perform significantly better or worse than the majority. In this case, we pursued an inhibitor that performed significantly better at inhibiting the *Plasmodium* interaction while not interfering with the human interaction. We compared both responses normalized to their respective controls, with human plotted on the *x*-axis and *P. falciparum* on the *y*-axis. Each point corresponds to the average response in the presence of one compound, with error bars corresponding to the standard deviation in the individual measurements. The black rectangle in the graph encloses responses that are within one standard deviation of the average non-control response.

Thermal shift assay. Small molecules identified from our VLS approach were tested in thermal stability assays in 1×PBS with 1:1800 final dilution of SYPRO orange dye (Invitrogen), with 65 μM His_6-*Pf*Atg8^{CM}. Triplicate measurements of fluorescence were measured from 20 to 80°C in a Bio-Rad C1000 thermal cycler for each concentration. ALC25 was added at increasing concentrations ranging from 100 μM to 2 mM, with DMSO concentration kept constant in all samples.

ALC25 binding site verification by SPR peptide-protein interaction assay. The peptide sequence NDWLLPSY of *Pf*Atg3

(PlasmoDB gene ID: *PF*3D7_0905700.1), as co-crystallized in complex with *Pf*Atg8 (PlasmoDB gene ID: *PF*3D7_1019900), PDB ID: 4EOY,[81] was synthesized and cloned into a pRSF-1b plasmid yielding Avitag-*Pf*Atg3 as previously described [86]. Avitag-*Pf*Atg3 was expressed in TB media and induced with 1 mM IPTG for 16 h at 20°C. Cells were lysed using an Avestin cell disruptor. The cell lysate was passed over TALON CellThru resin functionalized with Co^{2+} to capture the N-terminal His_{12} tag of the recombinant Avitag-*Pf*Atg3 construct. The eluted samples were validated by SDS-PAGE and directly used for capturing to a NeutrAvidin-coupled CM5 chip. A single lysine residue on the Avitag is biotinylated and forms a tight interaction with the NeutrAvidin on the chip. All interaction studies with the Avitag-*Pf*Atg3 CM5 chip were performed in 10 mM HEPES (pH 7.5), 150 mM NaCl, 3 mM EDTA, 1% DMSO.

ITC binding assay with ALC25. *Pf*Atg8^{CM} and *Pf*Atg8$^{CM/LD}$ were purified as previously described, and then dialyzed into a solution of 1× PBS overnight. ALC25 (100 µM) was titrated 10 µL at a time into a solution (1.8 mL) of the appropriate protein (10 µM) with a MicroCal VP-ITC machine (Malvern Instruments Inc., Westborough, MA, USA). The injections were 20 s in duration and spaced by 2 min. The jacket temperature was maintained at 25°C throughout. Data were analyzed and fit using Origin software (version 7.0, OriginLab, Northampton, MA, USA).

***Plasmodium* parasite growth inhibition assay**. *P. falciparum* 3D7 parasites were cultured as previously described [83]. Briefly, the parasites were cultured in 10% *v*/*v* human serum with 10 µg/mL gentamycin and synchronized with 5% *w*/*v* sorbitol; 1.5% ring-stage parasites were added to 96-well plates (Costar) containing serial dilutions, performed in triplicate, of compound in RPMI 1640 media (Gibco). Plates were incubated at 37°C in 5% O_2, 5% CO_2, and 90% N_2 for 72 h. Plates were frozen, thawed, and incubated with 100 µL 2×SYBR green in lysis buffer (20 mM Tris pH 7.5, 5 mM EDTA, 0.008% Saponin, 0.08% Triton X-100) in the dark for 1 h. Fluorescence was measured with a plate reader (HTS 7000, PerkinElmer) at excitation/emission wavelengths of 485/535 nm.

Immunoblot analysis of *P. falciparum* FCR3 blood-stage parasites. Immunoblot was performed as per protocol previously described [83], with 100 µM ALC25 or equivalent DMSO treatment for 5 h. In analyzing the quantity of *Pf*Atg8 on the gel, intensity was normalized to total protein amount in the protein stain with ImageJ [120].

Detection of *P. falciparum* EEFs *in vitro*. HC-04 human hepatocyte cell line (ATCC, Manassas, VA, USA) was maintained in complete medium (IMDM containing 2.5% FCS, 100 U/mL penicillin, 100 µg/mL streptomycin, and 2 mM l-glutamine, all from Gibco Life Technologies, Grand Island, NY, USA). *P. falciparum* 3D7HT-GFP parasite strain [106] was propagated in the Parasitology Core facility of the Johns Hopkins Malaria Research Institute. In vitro infection of human hepatocytes, as well as detection and quantification of *P. falciparum* 3D7HT-GFP EEFs by flow cytometry, were performed as previously described [121].

Hepatocyte cytotoxicity assay. Hepatocyte cytotoxicity assays were carried out in triplicate at two concentrations of inhibitory compounds using HC-04 cells as previously described [83]. Flow cytometry was used to assess two cytotoxicity markers by staining, namely propidium iodine (PI) and Annexin V.

Web servers and software. Amino acid sequences of apicomplexan Atg8 proteins were retrieved from EuPathDB [39] using the NCBI BLAST similarity search function [122] to find proteins homologous to *Pf*Atg8. Sequences of human Atg8 homologues were retrieved from UniProt [123]. Sequence alignments (Figure 2.1A,B) were performed using the ESPript multiple sequence alignment web server (http://espript.ibcp.fr) [124]. Clustal Omega was also used in sequence comparison (Figure 2.2C) [125]. FlowJo software was used in the analysis of flow cytometry data (Figure 2.8) [126]. The plot for Figure 2.4D was generated using Matlab [127], and all bar graphs and concentration curves were generated using GraphPad Prism [128]. Western blot quantification was performed with ImageJ (Figure 2.7B) [120]. Figure 2.3B was produced with OpenEye VIDA and rendered with POV-Ray [129], Figure 2.6A was produced with

OpenEye VIDA, and all other molecular graphics were created with PyMOL [130]. Table 2.1 structures were produced with ChemBioDraw [131].

Tables

Table 2.1: Small Molecule Probes of the Apicomplexan Loop Pocket

Comp.	IUPAC	CID	Structure	Mol. Wt.	cLogP[a]
ALC5	5-nitro-N-(pyridin-3-ylmethyl)furan-2-carboxamide	693996	[redacted]	247.21	0.68
ALC11	N-{[7-(pyridin-4-yl)-2,3-dihydro-1-benzofuran-2-yl]methyl}-4,5,6,7-tetrahydro-1-benzothiophene-3-carboxamide	45176114	[redacted]	390.50	4.81
ALC13	1-(4-bromobenzenesulfonyl)-N-(furan-2-ylmethyl)piperidine-4-carboxamide	1327632	[redacted]	427.31	2.19
ALC16	N-[(1-ethyl-1H-pyrazol-4-yl)methyl]-5-phenyl-1H-pyrrole-2-carboxamide	56751948	[redacted]	294.35	2.77
ALC17	N-[2-({6-ethyl-1-methyl-1H-pyrazolo[3,4-d]pyrimidin-4-yl}amino)ethyl]pyridin-3-amine	56894679	[redacted]	297.36	0.06
ALC18	N-[2-(N-methylmethanesulfonamido)ethyl]-5-phenyl-1H-pyrrole-2-carboxamide	56745406	[redacted]	321.40	2.12
ALC19	3-ethyl-N-{[2-(pyrazin-2-yl)-1,3-thiazol-4-yl]methyl}-1H-indole-2-carboxamide	50956056	[redacted]	363.44	2.38
ALC20	5-(2-chlorophenyl)-2-methyl-N-[(6-oxo-3,6-dihydropyrimidin-4-yl)methyl]furan-3-carboxamide	50965334	[redacted]	343.76	2.93
ALC22	2-ethyl-4-methyl-N-{[2-(pyrimidin-2-yl)-1,3-thiazol-4-yl]methyl}-1,3-thiazole-5-carboxamide	50974374	[redacted]	345.44	1.40
ALC24	N-{[1-(1H-imidazol-5-ylmethyl)pyrrolidin-3-yl]methyl}-8-methylimidazo[1,2-a]pyridine-2-carboxamide	56723508	[redacted]	338.41	0.35
ALC25	1-[(3-methylphenyl)methyl]-3-{[1-(pyridin-2-ylmethyl)pyrrolidin-3-yl]methyl}urea	56750287	[redacted]	338.45	1.59
ALC28	N-{[2-(dimethylamino)pyridin-3-yl]methyl}-4-phenyl-1H-imidazole-5-carboxamide	56749897	[redacted]	321.38	3.10
ALC29	N-{[3-(2-pyridinyl)-1,2,4-oxadiazol-5-yl]methyl}benzamide	28854470	[redacted]	280.3	1.09
ALC30	N-{[2-(pyrazin-2-yl)-1,3-thiazol-4-yl]methyl}-4H-thieno[3,2-b]pyrrole-5-carboxamide	50952642	[redacted]	341.41	1.71

[a] cLogP values calculated by researchers at ChemBridge Corporation

Figures

Figure 2.1. Atg8, a pan-apicomplexan drug target. (A) Atg8 sequence alignment for ten species of *Plasmodium* (ESPript, global score set to 0.9). *P. cynomolgi* has a 16-residue non-conserved N-terminal extension not shown here. The sequence is color-coded to indicate residues involved in ligand interaction sites: magenta, green, and orange correspond to the W-, L-, and A-sites, respectively. Ovals indicate side chain interactions, and triangles indicate backbone interactions involved in each site. The yellow box encloses the A-loop. (B) Sequence alignment of the A-loop in several apicomplexan species and human homologues (ESPript, global score set to 0.4). Note that the *N. caninum* and *T. gondii* Atg8 amino acid sequences are identical. The A-loop is notably absent from the human homologues.

Figure 2.1

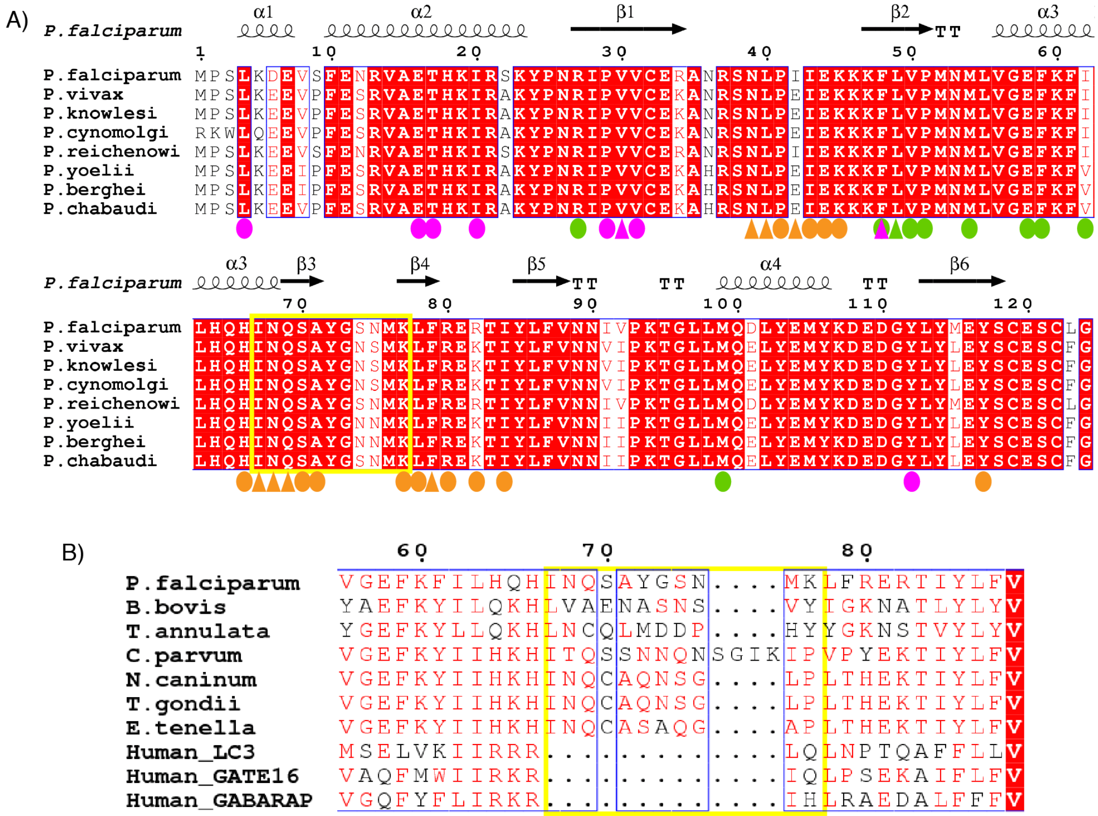

A)
P.falciparum
α1 α2 β1 β2 TT α3
1 10 20 30 40 50 60
P.falciparum MPSLKDEVSFENRVAETHKIRSKYPNRIPVVCERANRSNLPIIEKKKFLVPMNMLVGEFKFI
P.vivax MPSLKEEVPFESRVAETHKIRAKYPNRIPVVCEKANRSNLPEIEKKKFLVPMNMLVGEFKFI
P.knowlesi MPSLKEEVPFESRVAETHKIRAKYPNRIPVVCEKANRSNLPEIEKKKFLVPMNMLVGEFKFI
P.cynomolgi RKWLQEEVPFESRVAETHKIRAKYPNRIPVVCEKANRSNLPEIEKKKFLVPMNMLVGEFKFI
P.reichenowi MPSLKDEVSFENRVAETHKIRSKYPNRIPVVCERANRSNLPIIEKKKFLVPMNMLVGEFKFI
P.yoelii MPSLKEEIPFESRVAETHKIRAKYPNRIPVVCEKAHRSNLPEIEKKKFLVPMNMLVGEFKFV
P.berghei MPSLKEEIPFESRVAETHKIRAKYPNRIPVVCEKAHRSNLPEIEKKKFLVPMNMLVGEFKFV
P.chabaudi MPSLKEEVPFESRVAETHKIRAKYPNRIPVVCEKAHRSNLPEIEKKKFLVPMNMLVGEFKFV
P.falciparum
α3 β3 β4 β5 TT TT α4 TT β6
70 80 90 100 110 120
P.falciparum LHQHINQSAYGSNMKLFRERTIYLFVNNIVPKTGLLMQDLYEMYKDEDGYLYMEYSCESCLG
P.vivax LHQHINQSAYGNSMKLFREKTIYLFVNNVIPKTGLLMQELYEMYKDEDGYLYLEYSCESCFG
P.knowlesi LHQHINQSAYGNSMKLFREKTIYLFVNNVIPKTGLLMQELYEMYKDEDGYLYLEYSCESCFG
P.cynomolgi LHQHINQSAYGNSMKLFREKTIYLFVNNVIPKTGLLMQELYEMYKDEDGYLYLEYSCESCFG
P.reichenowi LHQHINQSAYGSNMKLFRERTIYLFVNNIVPKTGLLMQDLYEMYKDEDGYLYMEYSCESCLG
P.yoelii LHQHINQSAYGNNMKLFREKTIYLFVNNIIPKTGLLMQELYEMYKDEDGYLYLEYSCESCFG
P.berghei LHQHINQSAYGNNMKLFREKTIYLFVNNIIPKTGLLMQELYEMYKDEDGYLYLEYSCESCFG
P.chabaudi LHQHINQSAYGSNMKLFREKTIYLFVNNIIPKTGLLMQELYEMYKDEDGYLYLEYSCESCFG
B)
60 70 80
P.falciparum VGEFKFILHQHINQSAYGSN....MKLFRERTIYLFV
B.bovis YAEFKYILQKHLVAENASNS....VYIGKNATLYLYV
T.annulata YGEFKYLLQKHLNCQLMDDP....HYYGKNSTVYLYV
C.parvum VGEFKYIIHKHITQSSNNQNSGIKIPVPYEKTIYLFV
N.caninum VGEFKYIIHKHINQCAQNSG....LPLTHEKTIYLFV
T.gondii VGEFKYIIHKHINQCAQNSG....LPLTHEKTIYLFV
E.tenella VGEFKYIIHKHINQCASAQG....APLTHEKTIYLFV
Human_LC3 MSELVKIIRR.............LQLNPTQAFFLLV
Human_GATE16 VAQFMWIIRKR.............IQLPSEKAIFLFV
Human_GABARAP VGQFYFLIRKR.............IHLRAEDALFFFV

Figure 2.2 The apicomplexan loop site. (A) Structure of *Pf*Atg8 (PDB ID: 4EOY) color coded as in Figure 2.1A, showing the ligand interaction sites. B) Comparison of *Pf*Atg8 (PDB ID: 4EOY) and hLC3 (PDB ID: 2ZJD) structures, illustrating A-loop divergence. The yellow loop corresponds to the yellow box delineated in Figure 2.1. (C) A-loop pocket of *Pf*Atg8 color-coded based on conservation with five apicomplexan homologues. Green: identical or backbone interaction; yellow: conserved mutation; red: different. The orientation is the same as shown in panel (B).

Figure 2.2

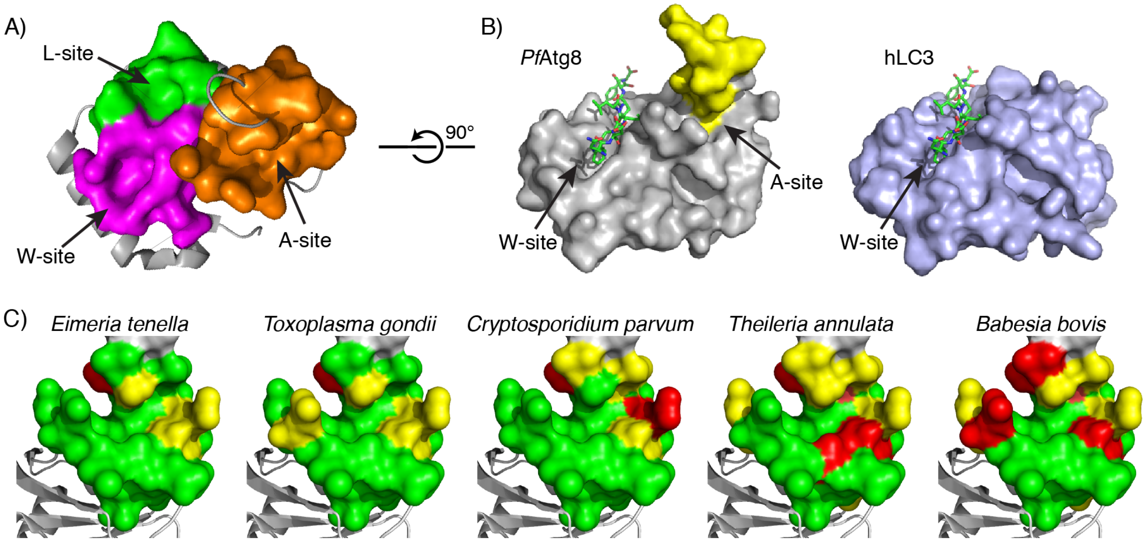

A)
L-site
W-site
A-site
90°
B)
PfAtg8
A-site
W-site
hLC3
W-site
C)
Eimeria tenella
Toxoplasma gondii
Cryptosporidium parvum
Theileria annulata
Babesia bovis

Figure 2.3 Generation of an *in silico Pf*Atg8 receptor for docking studies. (A) *Pf*Atg8 model used in OpenEye docking, with docking restraints indicated by green and magenta spheres. Side chain interaction: green; backbone interaction: magenta. (B) A-loop pocket size characteristics. The left image is oriented as in panel (A). (C) VLS flow diagram and attrition rate at each step.

Figure 2.3

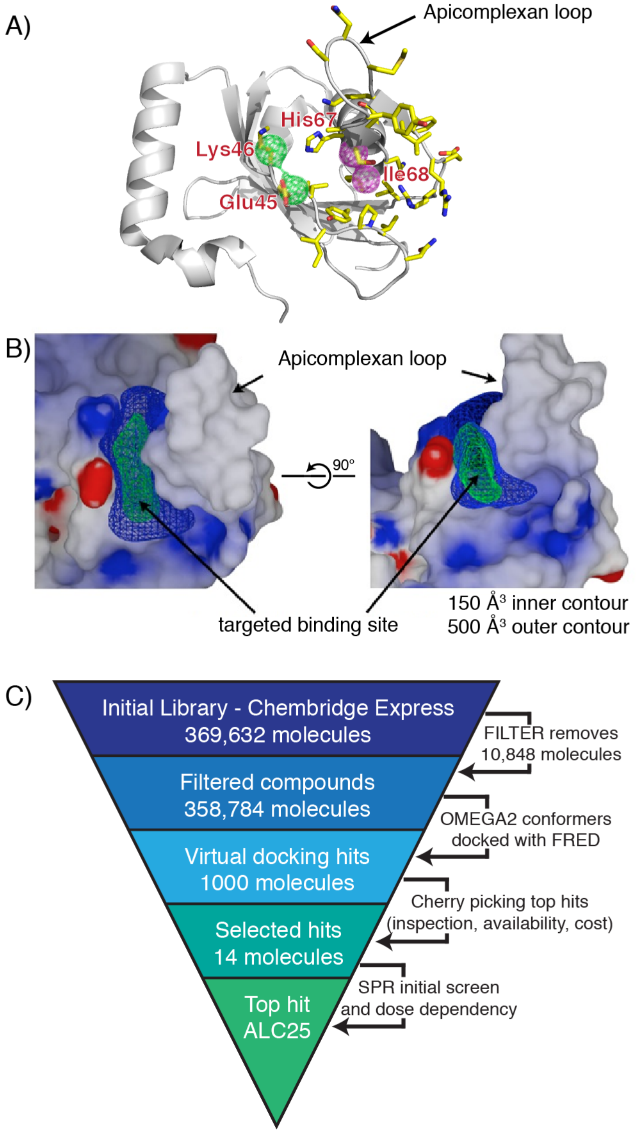

A)
Apicomplexan loop
His67
Lys46
Ile68
Glu45
B)
Apicomplexan loop
90°
targeted binding site
150 Å³ inner contour
500 Å³ outer contour
C)
Initial Library - Chembridge Express
369,632 molecules
FILTER removes
10,848 molecules
Filtered compounds
358,784 molecules
OMEGA2 conformers
docked with FRED
Virtual docking hits
1000 molecules
Cherry picking top hits
(inspection, availability, cost)
Selected hits
14 molecules
SPR initial screen
and dose dependency
Top hit
ALC25

Figure 2.4 PPI inhibition assayed by SPR. All binding responses are normalized to the baseline control at 100%. (A) Schematic of SPR interaction study. The two images illustrate SPR interactions in the absence (left) and presence (right) of compound. (B) Effect of compounds on *Pf*Atg8-*Pf*Atg3 interaction. The green dotted line corresponds to control response. Blue indicates some precipitation of protein or compound observed at 500 μM concentration. (C) Effect of compounds on hAtg3-hLC3 interaction. Green: control; blue: observed precipitation. (D) The TSIS plot for human and *P. falciparum* data. Human (*x*-axis) versus *Plasmodium* (*y*-axis) normalized SPR binding. Dotted lines are the control response. The gray rectangle encloses compounds within one standard deviation of average inhibited response. The green rectangle encloses favorable area for potential antimalarial molecules.

Figure 2.4

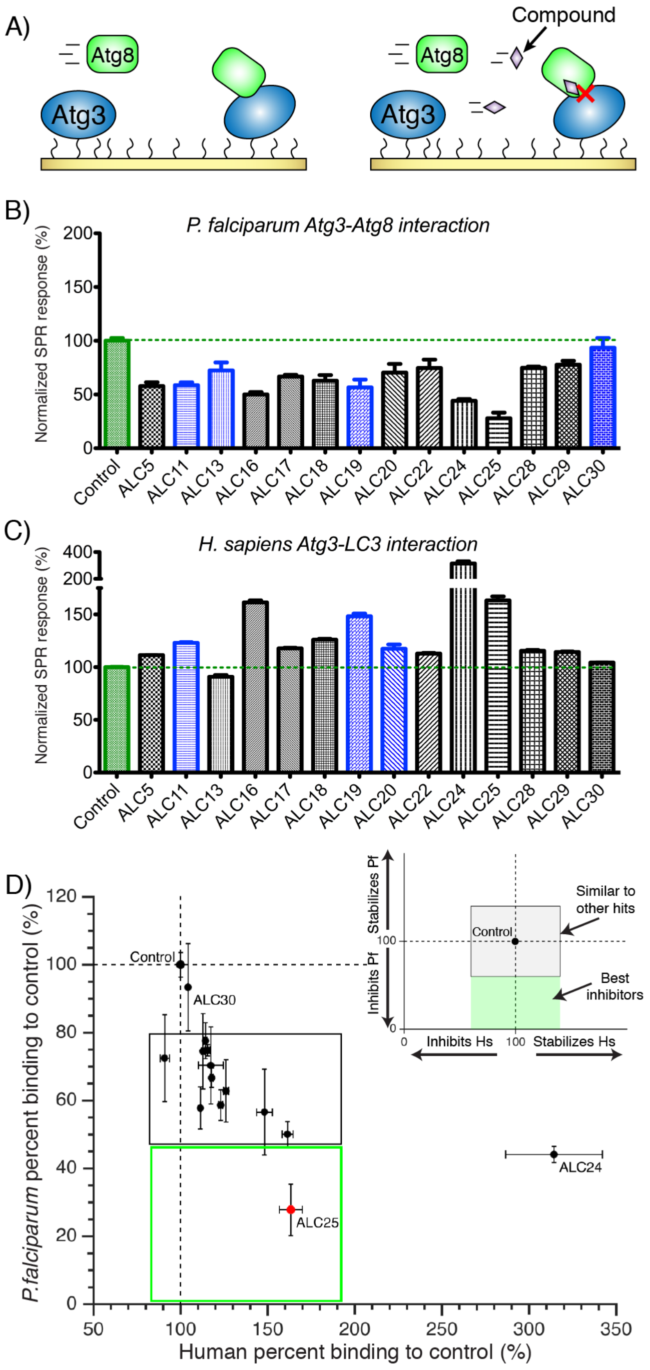

A)
Atg8
Atg3
Compound
Atg8
Atg3
B)
P. falciparum Atg3-Atg8 interaction
Normalized SPR response (%)
200
150
100
50
0
Control
ALC5
ALC11
ALC13
ALC16
ALC17
ALC18
ALC19
ALC20
ALC22
ALC24
ALC25
ALC28
ALC29
ALC30
C)
H. sapiens Atg3-LC3 interaction
Normalized SPR response (%)
400
200
150
100
50
0
Control
ALC5
ALC11
ALC13
ALC16
ALC17
ALC18
ALC19
ALC20
ALC22
ALC24
ALC25
ALC28
ALC29
ALC30
D)
P.falciparum percent binding to control (%)
120
100
80
60
40
20
0
Control
ALC30
ALC24
ALC25
50
100
150
200
250
300
350
Human percent binding to control (%)
Stabilizes Pf
Inhibits Pf
100
0
Control
Similar to other hits
Best inhibitors
0
100
Inhibits Hs
Stabilizes Hs

Figure 2.5 ALC25 characterization. (A) IC_{50} of ALC25 in SPR binding assay. (B) *Pf*Atg8 melting temperature in the presence of increasing concentrations of ALC25. (C) SPR binding response of *Pf*Atg8 binding to Avitag-*Pf*Atg3 peptide in the presence and absence of ALC25. Binding of *Pf*Atg8 to the peptide, which binds to the W/L site, is not inhibited by ALC25. Error bars represent ±SD of triplicate experiments.

Figure 2.5

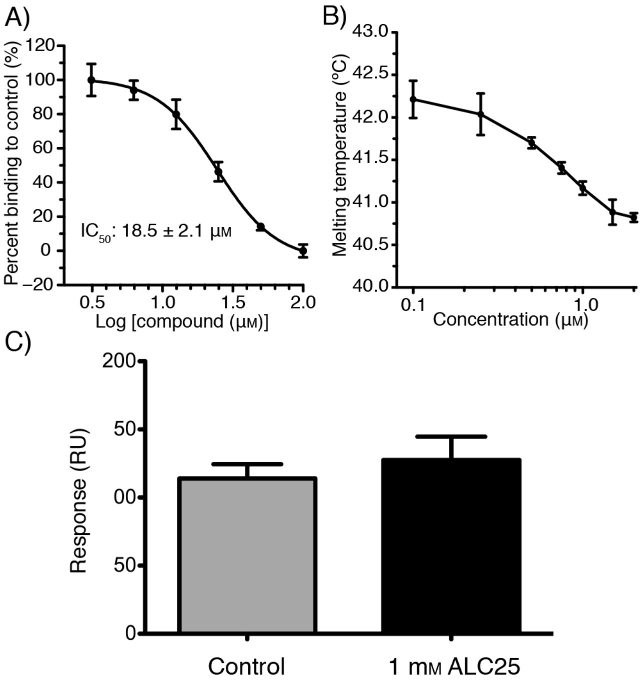

A)
Percent binding to control (%)
120
100
80
60
40
20
0
−20
0.5
1.0
1.5
2.0
Log [compound (μM)]
IC$_{50}$: 18.5 ± 2.1 μM
B)
Melting temperature (°C)
43.0
42.5
42.0
41.5
41.0
40.5
40.0
0.1
1.0
Concentration (μM)
C)
Response (RU)
200
50
00
50
0
Control
1 mM ALC25

Figure 2.6 ALC25 FRED docking report. (A) ALC25 docking pose showing four hydrogen bond interactions with *Pf*Atg8. (B) Schematic representation of the ligand environment and pocket characteristics, showing that most of the ALC25 ligand occupies the A-loop pocket. A region near the pyrrolidine ring would allow further extension of the ligand. (C) ALC25 had the second-highest docking score of all tested compounds in the filtered ChemBridge core library. (D) The residue fingerprint highlights residues in proximity to the ligand; residues in bold are in contact with the ligand, and are further distinguished by color based on the type of interaction. (E) Additional properties for ligand optimization guidance are given based on shape, hydrogen bonding interaction, and protein and ligand desolvation energies.

Figure 2.6

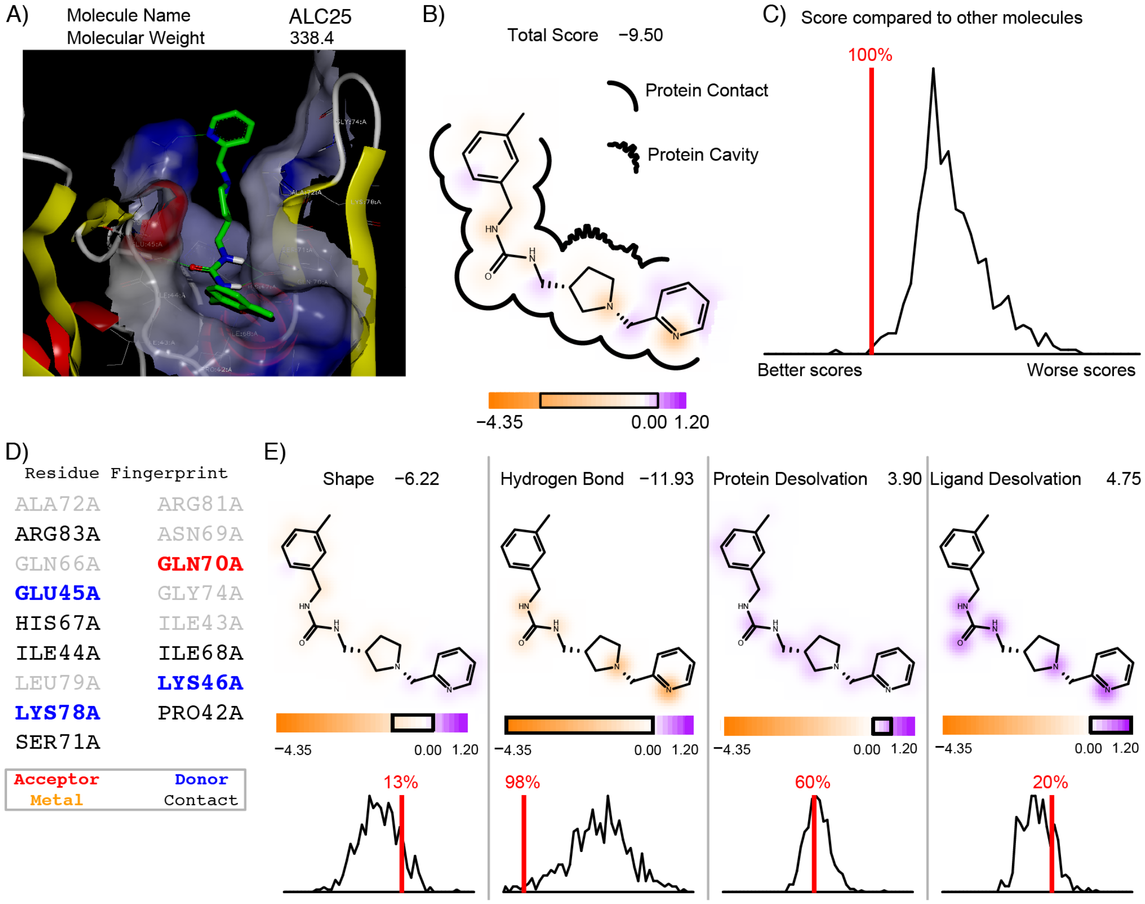

A)
Molecule Name
ALC25
Molecular Weight
338.4
B)
Total Score -9.50
Protein Contact
Protein Cavity
HN
H N
O
N
N
−4.35
0.00 1.20
C) Score compared to other molecules
100%
Better scores
Worse scores
D)
Residue Fingerprint
ALA72A
ARG81A
ARG83A
ASN69A
GLN66A
GLN70A
GLU45A
GLY74A
HIS67A
ILE43A
ILE44A
ILE68A
LEU79A
LYS46A
LYS78A
PRO42A
SER71A
Acceptor
Donor
Metal
Contact
E)
Shape -6.22
Hydrogen Bond -11.93
Protein Desolvation 3.90
Ligand Desolvation 4.75
−4.35
0.00
1.20
13%
98%
60%
20%

Figure 2.7 Effect of ALC25 on blood-stage parasite development.

(A) IC_{50} of ALC25 and CQ in blood-stage parasites as determined through the SYBR green I assay. (B) Western blot: the top section of the western blot shows the protein stain, indicating apparent uniform protein loading across the lanes. The lower molecular weight range of the gel was probed with anti-*Tg*Atg8 antibody, which cross-reacts with *Pf*Atg8. Treatment with 100 μM ALC25 for 5 h led to a shift in *Pf*Atg8 mobility, corresponding to a shift from a predominantly lipidated (black dotted line) to unlipidated (gray dotted line) population. (C) Quantification of *Pf*Atg8 on western blot in panel (B).

Figure 2.7

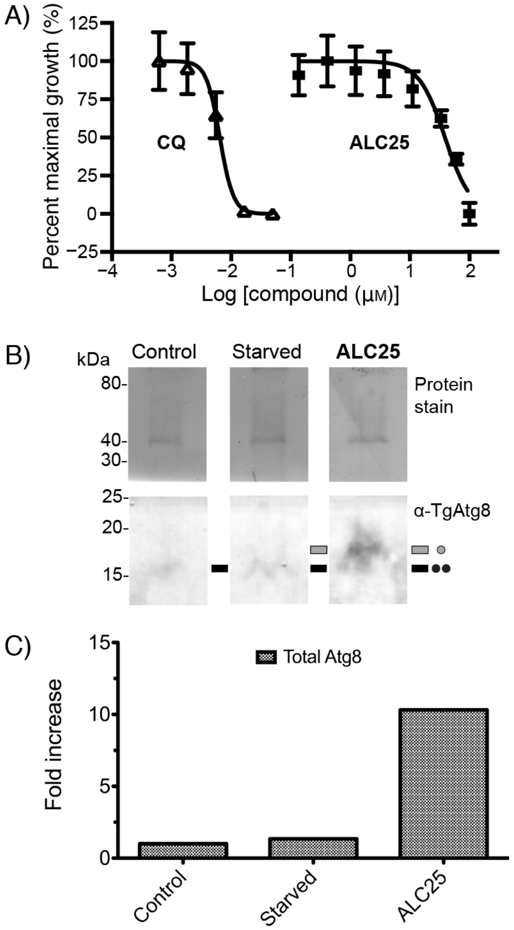

A)
Percent maximal growth (%)
125
100
75
50
25
0
−25
−4
−3
−2
−1
0
1
2
Log [compound (μM)]
CQ
ALC25
B)
kDa
Control
Starved
ALC25
80-
40-
30-
Protein stain
25-
20-
15-
α-TgAtg8
C)
Fold increase
15
10
5
0
Total Atg8
Control
Starved
ALC25

Figure 2.8 Effect of ALC25 on liver-stage parasite development.

(A) Flow-cytometry-based detection of *P. falciparum* 3D7HT-GFP EEFs propagated in the presence of ALC25 or DMSO as a vehicle control until 96 h post-infection. Representative dot plots for each condition are shown. Numbers reflect percentages of viable infected cells (GFP+/PI-). The bar graph gives the mean (±SD) of EEF percentages from three independent hepatocyte cultures. Non-infected and DMSO controls previously published [83]. (B) Representative dot plots illustrating percentages of viable cells (Annexin V-/PI-) detected by flow cytometry in HC-04 cultures treated with ALC25 or DMSO (vehicle control) for 96 h. The bar graph gives the mean (±SD) for three independent hepatocyte cultures.

Figure 2.8

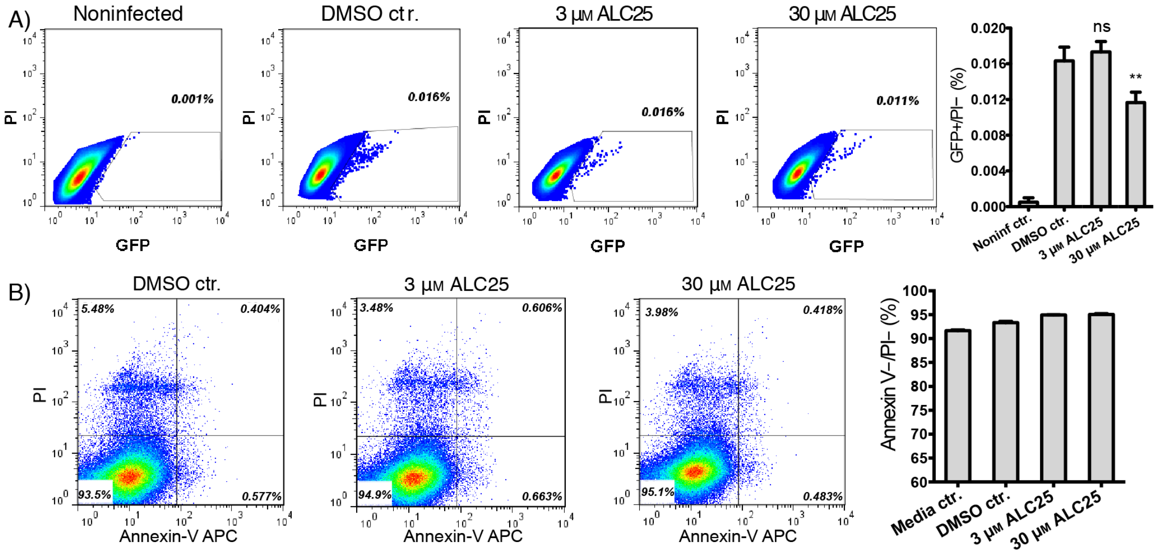

A)
Noninfected
DMSO ctr.
3 μM ALC25
30 μM ALC25
PI
GFP
0.001%
0.016%
0.016%
0.011%
GFP+/PI− (%)
0.000
0.004
0.008
0.012
0.016
0.020
ns
**
Noninf ctr.
DMSO ctr.
3 μM ALC25
30 μM ALC25
B)
DMSO ctr.
3 μM ALC25
30 μM ALC25
PI
Annexin-V APC
5.48%
0.404%
93.5%
0.577%
3.48%
0.606%
94.9%
0.663%
3.98%
0.418%
95.1%
0.483%
Annexin V−/PI− (%)
60
65
70
75
80
85
90
95
100
Media ctr.
DMSO ctr.
3 μM ALC25
30 μM ALC25

Figure S2.1. ChemBridge Core Library properties. (A) The cLogP distribution; (B) the molecular weight distribution; (C) the percentage of chiral molecules in the library; and (D) an Omega2 conformer plot highlighting a significant amount of flexible molecules present in the library.

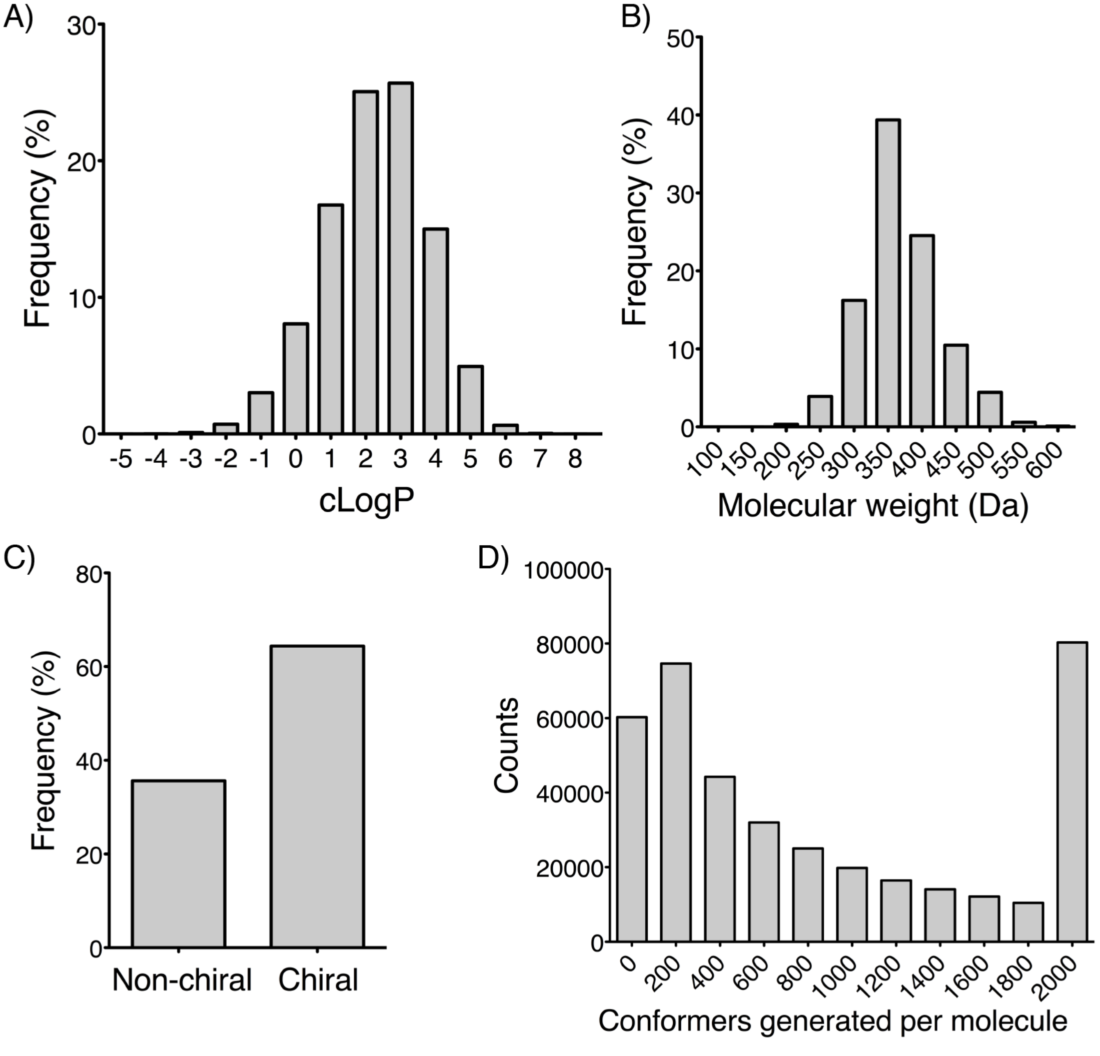
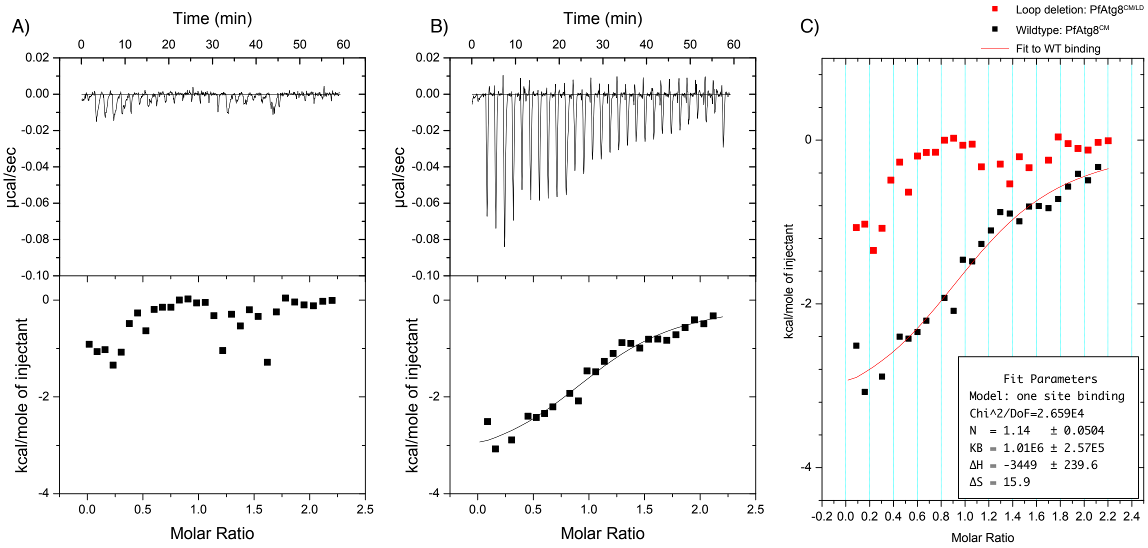


Figure S2.2. ITC with ALC25. (A) The raw ITC curve and the ΔH from integration for a titration of 50 μM ALC25 into 5 μM *Pf*Atg8$^{\mathrm{CM/LD}}$. No observable heat of binding between ALC25 and the loop deletion mutant. (B) The raw ITC curve and the ΔH from integration for a titration of 50 μM ALC25 into 5 μM *Pf*Atg8$^{\mathrm{CM}}$. Fits to a one-site binding model. (C) Both graphs of ΔH are placed on the same axes for ease of comparison. The fit parameters for ALC25 binding to *Pf*Atg8$^{\mathrm{CM}}$ are also included.



CHAPTER 3

Testing Compounds Against the apicomplexan Atg8–Atg3 Interaction

Acknowledgements

The following chapter represents predominantly unpublished work. However, as the third author on [85], I helped to detect and quantify the presence of lipidated and unlipidated *Pf*Atg8 in PTA5-treated, infected cells via Western Blot. These images have been reproduced below with permission.

I acknowledge Dr. Hain for her work in the initial screening of the MMV library against the *Pf*Atg3-*Pf*Atg8 interaction and subsequent *in vitro* and *in vivo* measurements, as published in [83]. I also thank the laboratory of Dr. Caren Freel-Meyers for the synthesis of PTA4, PTA5, and the PTA derivative library. Dr. J. Smith from Seattle BioMed provided the FCR3-CSA strain of parasite used in the flow cytometry assay as well as the *Pf*EMP1 antibody used in that assay. Dr. J. Levitskaya provided the isotype control antibody and valuable flow cytometry advice.

Abstract

The Medicines for Malaria Venture (MMV) Malaria Box of 400 compounds was tested against the *P. falciparum* Atg8–Atg3 interaction via Surface Plasmon Resonance (SPR), leading to the identification of three inhibitory compounds that shared a common scaffold: 4-pyridin-2-yl-1,3-thiazol-2-amine (PTA). One of these compounds, PTA1, was tested in parasite cultures: it had an *in vivo* IC_{50} of 1.6 μM and led to a marked increase in the level of Atg8 and a shift to a predominantly unlipidated population as assayed via Western Blot. Based on these promising results, a PTA-derivative library was generated, consisting of 240 unique molecules sharing the PTA scaffold. In this work, the library of compounds is tested against multiple homologous Atg8–Atg3 interactions from the following species: *Plasmodium falciparum*, *Toxoplasma gondii*/*Neospora caninum*, *Eimeria tenella*, and *Cryptosporidium parvum*. This derivative library screening yields numerous potential hits. Using the same SPR setup and the same protein homologue pairs, the hits from the Virtual Ligand Screening against *Pf*Atg8 are also screened. Finally, several PTA-derivative molecules are tested and found to decrease infected red blood cell (iRBC) surface-exposed *P. falciparum* erythrocyte membrane protein 1 (*Pf*EMP1) levels. These results are promising for future antimalarial and anti-apicomplexan drug development.

Introduction

The following work is an effort to explore the effects on apicomplexan Atg3-Atg8 interactions of two groups of molecules: (1) a class of 4-pyridin-2-yl-1,3-thiazol-2-amine (PTA) derivative drugs and (2) virtual ligand screening (VLS) hits, both shown to have activity against the *Pf*Atg3-*Pf*Atg8 interaction [83,85].

As discussed in previous chapters, the autophagy pathway or, more specifically, the Atg8-conjugation pathway within apicomplexans is a strong candidate for drug development, based on its essentiality within investigated species (*P. falciparum* and *T. gondii*). The value in targeting Atg8 in multiple species simultaneously is that, while each protein target is unique, such an approach takes advantage of common features and high sequence similarity to potentially identify novel classes of anti-apicomplexan molecules.

The identification of a novel antimalarial, in particular, is important not only because of increased drug resistance, but also because current medications are not as effective in treating severe malaria. Cerebral malaria is one manifestation of severe malaria, characterized by impaired consciousness. For children, a diagnosis of cerebral malaria is a death sentence without treatment; even with it, 20% will not survive, and many survivors will suffer long-term neurological damage [132]. This form of malaria is linked to the expression of *Plasmodium falciparum* erythrocyte membrane protein 1 (*Pf*EMP1) [133].

*Pf*EMP1 is involved in evading host immunity: the protein is exported from the parasitophorous vacuole to knobs on the surface of infected red blood cells (iRBCs). *Pf*EMP1 then binds to receptors on various host cells, sequestering the infected cells from circulation and allowing the iRBCs to avoid clearance by the spleen [134,135]. Because *Pf*EMP1 is a major virulence factor, it has been the target of numerous studies, and has become an important malaria vaccine candidate [136]. Direct targeting of *Pf*EMP1 is complicated by the fact that the malaria parasite expresses only one of approximately 60 *Pf*EMP1 variants at any given time [137]. An equally viable and perhaps more accessible approach would be to inhibit the export of *Pf*EMP1 to the iRBC membrane.

Along these lines, there is reason to believe that *Pf*Atg8 may be involved in export to the host cell. As mentioned in Chapter 1, *Pf*Atg8 has been found to colocalize with the Maurer's cleft (MC), an intermediate structure involved in the export pathway [71]. Furthermore, *Pf*Atg8 has been found to colocalize with *Pf*EMP1 and to partially colocalize with *P. falciparum* skeleton binding protein 1 (*Pf*SBP-1) [138]. *Pf*SBP-1 is involved in *Pf*EMP1 export and localizes to the MC [139]. While *Pf*SBP-1 is one of many proteins exported to the host, it does not contain the canonical *Plasmodium* export element (PEXEL) motif typical of other exported proteins, and *Pf*EMP1 is missing a hydrophobic signal sequence also characteristic of exported proteins, leaving open questions as to how these proteins make their way to the host. To test the involvement of *Pf*Atg8 with this process, we used two of the PTA inhibitors in a flow cytometry-based assay and probed for surface-exposed *Pf*EMP1 after treatment.

Results

Sequence similarity and homology modeling of PfAtg8 homologues confirm feasibility of drug repositioning

The sequence for Atg8 is well-conserved within the apicomplexan phylum. In particular, a multiple sequence alignment of *Plasmodium falciparum* Atg8 with the homologous proteins from *Cryptosporidium parvum*, *Eimeria tenella*, *Neospora caninum*, and *Toxoplasma gondii* shows greater than 60% sequence identity between *Pf*Atg8 and each of the other molecules (Figure 3.1A). Given that the structure of *Pf*Atg8 is known, this level of sequence identity makes these protein sequences good candidates for homology modeling. I-TASSER is a protein structure and function prediction server supported by the Zhang Lab at the University of Michigan [140]. It has received first place ranking in the last four Critical Assessment of Techniques for Protein Structure Prediction (CASP) events for homology modeling [141–144]. Using this server, the structures of *Cp*Atg8, *Et*Atg8, and *Nc*/*Tg*Atg8 were generated (Figure 3.1B). A comparison of these structures shows clear areas of conservation, both in the W- and L- sites responsible for binding to Atg3 and in the A-site unique to the apicomplexan Atg8 homologues (Figure 3.1C,D).

In the following studies, we take advantage of these areas of conservation by testing compounds targeted against the *Pf*Atg3-*Pf*Atg8 interaction against the Atg3-Atg8 interactions of other apicomplexan organisms.

VLS compounds tested against homologous Atg8-Atg3 interactions for C. parvum, E. tenella, and N. caninum via SPR PPI assay

As described in Chapter 2, we used our SPR assay to test the effect of the VLS compounds against the human and *P. falciparum* Atg3-Atg8 protein-protein interactions (Figure 2.4). Using a similar setup, we tested the same set of VLS compounds against the Atg3-Atg8 interactions for *C. parvum*, *E. tenella*, and *N. caninum/T. gondii*. The results for these three inhibition assays are shown in Figure 3.2.

Compounds 11, 13, and 22 resulted in a significant inhibition of the *Cp*Atg3-*Cp*Atg8 protein-protein interaction as compared to the control (Figure 3.2A). However, for all Atg8 homologues except *Plasmodium*, some precipitation of the protein Atg8 was observed upon addition of compound 13, undermining the apparent efficacy of this compound as an inhibitor.

Compounds 5, 29, and 30 resulted in significant inhibition of the *Et*Atg3-*Et*Atg8 interaction (Figure 3.2B), and in our SPR assay with

*Nc*Atg3 immobilized on the chip and *Nc/Tg*Atg8 flowing over it, compounds 5, 11, 13, 16 and 22 all resulted in a decreased response (Figure 3.2C).

PTA-derivative library created based on favorable initial findings

The Medicines for Malaria Venture (MMV) made freely available a Malaria Box of 200 drug-like and 200 probe-like molecules known to have activity against blood-stage *P. falciparum* malaria [82]. These compounds were previously screened against the *Pf*Atg3-*Pf*Atg8 interaction in a surface plasmon resonance (SPR) assay, with *Pf*Atg3 immobilized on a chip and *Pf*Atg8 flown over in the presence of each compound [83]. Comparing the inhibited response to the control response resulted in the identification of six molecules that produced greater than a 25% inhibited response. Of these, three molecules showed dose-dependent inhibition and shared the same PTA scaffold: N-(4-methylphenyl)-4-pyridin-2-yl-1,3-thiazol-2-amine (PTA1), 2-methylsulfanyl-N-(4-pyridin-2-yl-1,3-thiazol-2-yl)benzamide (PTA2), and 2-bromo-N-(4-pyridin-2-yl-1,3-thiazol-2-yl)benzamide (PTA3) (Table 1).

PTA1 had an IC_{50} of 18.16 ± 0.25 µM *in vitro* in SPR and 1.6 ± 1.5 µM *in vivo* in parasite culture as tested via a SYBR green I growth inhibition assay. Not only was it effective in killing the parasites, but it also produced a noticeable shift in *Pf*Atg8 lipidation state – the protein band shifted from a predominantly lipidated to a predominantly unlipidated population (Figure 3.3A). This shift was also accompanied by a remarkable upregulation of *Pf*Atg8 (Figure 3.3B). The shift in lipidation state was expected: inhibition of the *Pf*Atg3-*Pf*Atg8 interaction would inhibit the subsequent conjugation of *Pf*Atg8 to PE. As for the upregulation, it is possible that the parasite responded to the loss of lipidated Atg8 by significantly upregulating production of the protein.

Based on these results, a PTA-benzaldehyde derivative was synthesized with a functional handle for the generation of a library of PTA-derivative molecules: 4-formyl-N-(4-pyridin-2-yl-1,3-thiazol-2-yl)-benzamide (PTA4) (Table 1). An unreactive version of PTA4, 4-(Dimethoxymethyl)-N-(4-pyridin-2-yl-1,3-thiazol- 2-yl)benzamide (PTA5), was synthesized to test the inhibitory effects of this derivative (Table 1). PTA5 had an *in vitro* IC_{50} of 2.86 µM in SPR and *in vivo* IC_{50} of 1.48 ± 0.6 µM in a SYBR green assay.

Thanks to our collaborators in the lab of Dr. Caren Freel-Meyers, the PTA-derivative library of 232 molecules was created through tethering PTA to a library of aldehydes via a dialkoxyamine linker. The set of linkers used to create the library is shown in Figure 3.4.

PTA-derivative library tested against five different homologous Atg3-Atg8 interactions via SPR interaction assay

We used our established SPR interaction assay to test the inhibitory effects of the PTA derivatives on the Atg3-Atg8 interaction for human, *P. falciparum*, *C. parvum*, *E. tenella*, and *N. caninum*/*T. gondii* homologues. In each case, the Atg3 homologue was immobilized on the SPR chip, with Atg8 flown over in the presence or absence of PTA derivatives at 5 μM concentration. To analyze the data, we made use of TSIS plots (see Chapter 2), comparing the human response for each analyte on the x-axis to the apicomplexan response on the y-axis (Figure 3.5). We then identified inhibitors that produced a response more than two standard deviations lower than the average inhibited response.

For the *P. falciparum* interaction, the inhibited responses were so much lower than the control responses at 5 μM that the screen was repeated at 0.5 μM in an attempt to isolate the most effective of the inhibitors. In this second screen, the following analytes produced a significantly decreased SPR response: 42, 43, 44, 45, 46, 47, 48, 49, 123, 124, 125, 126 (Figure 3.5A).

For the *C. parvum* interaction, no analytes produced an SPR response more than two standard deviations below the average inhibited response (Figure 3.5B). Setting the threshold to 1.5 standard deviations yielded only three hits: 130, 156, and 159. Of the apicomplexan Atg8 homologues, *Cp*Atg8 was the most divergent, which could account for the lack of hits.

In the case of the *E. tenella* Atg3-Atg8 interaction, analytes 8, 46, 50, and 55 produced a significantly inhibited response (Figure 3.5C).

We were unable to successfully clone *Tg*Atg3 after multiple attempts. However, since *Nc*Atg8 and *Tg*Atg8 share 100 percent sequence identity, we were able to use *Nc*Atg3 as the binding partner for *Nc/Tg*Atg8. Presumably, inhibition of this interaction could have implications for both the *N. caninum* and the *T. gondii* interactions. In this case, for the *Nc*Atg3–*Nc/Tg*Atg8 interaction, analytes 41, 67, 72, 74, 75, 76, and 80 produced a significantly inhibited response (Figure 3.5D).

Pf*EMP1 expression on surface of iRBCs decreases in response to treatment with PTA1 or PTA5*

If *Pf*Atg8 is involved in exporting proteins to the host cell membrane, then interruption of *Pf*Atg8 lipidation may decrease the levels of exported proteins; in particular, we were interested in measuring the level of *Pf*EMP1 via flow cytometry. As previously mentioned, *Pf*EMP1 is a protein with approximately 60 variants encoded by the *var* multigene family, the expression of which is altered to enable the parasite to evade detection by the immune system. To be able to probe for *Pf*EMP1, we used the *P. falciparum* FCR3-CSA strain of parasite that had been selected for binding to chondroitin sulfate A (CSA), a placental receptor. Only one *Pf*EMP1 variant, encoded by the var2csa gene, binds to CSA, allowing for the use of specific antibodies raised against this variant.

Using blood infected with this strain, we treated late ring stage parasites with PTA1 or PTA5 at 3 μM and 30 μM concentrations for 16 hours. After 16 hours of treatment, we probed for *Pf*EMP1 with rabbit antisera raised against var2csa or with control rabbit antisera. The cells were then stained with a secondary antibody and with Thermo Fisher® DRAQ5™, a fluorescent DNA intercalator, to differentiate between infected and uninfected cells. One million cells were screened for each sample.

We removed clumped or dead cells from consideration by gating based on forward and side scatter (Figure 3.6A). To isolate the parasite-infected cells, we measured DRAQ5 fluorescence, gating conservatively to ensure the population under consideration predominantly consisted of iRBCs (Figure 3.6B). Based on DRAQ5+ staining, the percentage of iRBCs did not substantially differ between the control sample and the samples treated with 3 μM of either PTA1 or PTA5. When treated with 30 μM PTA1, however, the percentage of infected cells dropped to 50% and for 30 μM PTA5, it dropped to 73%.

We compared *Pf*EMP1 staining in the treated samples to the *Pf*EMP1-positive population in the DMSO control. In the 3 μM PTA1 treatment, the percent of iRBCs positively stained for *Pf*EMP1 dropped to 87.6% and down to 40.3% for 30 μM PTA1 treatment. Treatment with 3 μM PTA5 did not lead to a decrease in the *Pf*EMP1-positive population, but 30 μM PTA5 led to 58.8% *Pf*EMP1-positive staining.

Discussion and Future Directions

The high degree of sequence similarity between the apicomplexan Atg8 homologues (Figure 3.1) makes it possible to consider screening similar inhibitors against the Atg3-Atg8 protein-protein interaction (PPI) for multiple organisms. Furthermore, such an approach has the benefit of streamlining further drug development efforts. Identifying a class of molecules with inhibitory effects could substantially decrease future costs associated with the testing and development of this class.

From the set of apicomplexan loop compounds (ALCs), in each case, several showed significant inhibition of the Atg3-Atg8 PPI (Figure 3.5). As with the verification of ALC25, the next step would be to investigate SPR dose-dependency and then begin *in vivo* assays. There was some difficulty in dose-dependent studies as the ALCs showed only limited inhibition and poor solubility, making it difficult to explore a range of concentration values. It would be informative in each case to take several of the compounds and attempt other *in vitro* dose dependency assays or simple *in vivo* assays. In general, given that these compounds were identified in an *in silico* screen against the A-loop of *Pf*Atg8, it is not surprising that they exhibit weak inhibition of other PPIs.

In the case of the PTA-derivative molecules, the results were more encouraging. Excluding the *C. parvum* interaction, in each case, there were a number of PTA derivatives that produced an inhibited response more than two standard deviations below the average inhibited response.

This improved inhibitory success rate was to be expected based on the higher degree of conservation in the apicomplexan Atg8 W- and L- sites. These are the primary interaction sites for Atg3, and the PTA molecules are thought to bind here [83]. The variety of linkers featured in the library allowed for tailored binding to each molecule.

The data featured here represents only a preliminary screen of the molecules, however. Given more time, the next steps would involve many of the assays featured in Chapter 2 and in [83]. In vitro assays would include testing for dose dependency via SPR and direct binding via ITC. Of course, in each case, crystal structures and co-crystals with bound ligand would be informative regarding the binding site and mode of inhibition. Assuming further *in vitro* assays produced encouraging results, *in vivo* testing would be the next step. As in the case for PTA1 and PTA5, we would expect to see significant inhibition of parasite growth with treatment. As is the case for the plasmodial homologue, *Et*Atg8, *Nc*Atg8, and *Tg*Atg8 are synthesized with an exposed C-terminal glycine, meaning that they also likely exist in a predominantly lipidated state. Given this fact, we would expect to see similar changes in the lipidation states of these proteins upon treatment with PTA derivatives.

Finally, taking a closer look at the effect of the PTA molecules on *P. falciparum* in particular, the results of the *Pf*EMP1 inhibition assay are promising. Treatment with PTA1 and PTA5, which have been shown to inhibit the interaction between *Pf*Atg8 and *Pf*Atg3, led to a substantial decrease in the level of *Pf*EMP1-positive staining on the surface of iRBCs. Based only on these results, it is impossible to know the mode by which PTA1 and PTA5 led to *Pf*EMP1 export inhibition, but it is possible that this inhibition is the result of the change in *Pf*Atg8 lipidation state.

To determine the veracity of this claim, in-depth follow up would be required. First and foremost, flow cytometry with fixed and permeabilized cells would allow us to test the overall levels of *Pf*EMP1 in infected cells. These overall levels could also be probed via western blot analysis. If the export pathway is being affected, one would not necessarily expect to see a decrease in the overall level of *Pf*EMP1 expression, only in the amount that successfully reached the surface of the cell. Immunofluorescence analysis of treated cells would also allow us to visually monitor a change in *Pf*EMP1 localization, though this technique may not be large-scale enough to definitively detect a difference between untreated and treated cells [138].

Materials and Methods

Multiple sequence alignment and homology modeling. The sequences for the proteins in this study were downloaded from EuPathDB: *P. falciparum* Atg8 (EuPathDB gene ID: PF3D7_1019900), *C. parvum* Atg8 (cgd7_3990), *E. tenella* Atg8 (ETH_00016760), *N. caninum*/*T. gondii* Atg8 (NCLIV_008410 and TGME49_254120) [39]. The multiple sequence alignment was performed and rendered by Clustal

Omega [125]. The homology models for *Cp*Atg8, *Et*Atg8, and *Nc*/*Tg*Atg8 were generated using the online I-TASSER server [140]. Each sequence was submitted with default parameters and the top scoring model was chosen as a representative homology model for each protein. Figure 3.1 panels B-D were rendered using PyMOL [130] and panels C and D were colored based on the color_by_conservation.py script [145].

Plasmid and strain construction. The plasmid and strain construction for *Pf*Atg8, *Pf*Atg3, hLC3, and hAtg3 is as in [81] and [85]. The sequences for *Cp*Atg3, *Cp*Atg8, *Et*Atg3, and *Et*Atg8 were determined by homology to *Pf*Atg3 or *Pf*Atg8 using the online database EuPathDB [39], and then confirmed by sequence homology from BLAST searches. These genes were codon-optimized for expression in *E. coli* and ordered as gBlocks® gene fragments from IDT®, with encoded *Nc*oI and BamHI restriction sites at the beginning and end of the gene. The gene blocks were digested with New England BioLabs® (NEB) *Nc*oI and BamHI restriction enzymes and ligated with NEB T4 DNA ligase into a modified pRSF vector containing an N-terminal maltose binding protein (MBP) tag followed by a tobacco etch virus (TEV) protease cleavage site [86]. These plasmids were transformed into chemically competent *Escherichia coli* BL21 strain cells.

We received the *Tg*Atg8 construct from Anthony Sinai at the University of Kentucky College of Medicine. The sequence was PCR amplified with primers 5'-CACC ATGC CATC GATT CGCG ACGA AGTG

TCC (sense) and 5'-ttac ccca gagt gttc tctg aaga gtat tcca cgta ca (antisense). We received *N. caninum* cDNA from Wesley C. Van Voorhis at the University of Washington, Seattle. We retrieved the *Nc*Atg3 sequence from EuPathDB [39] (gene id: NCLIV_050410) to design primers, 5'-CATG CCAT GGCG CTTC CTCA AAAC G (sense) and 5'-cggg atcc cgct actt ctct ttcg tgtg gt (antisense), which we used to PCR amplify the sequence from the cDNA. The PCR products for *Tg*Atg8 and *Nc*Atg3 were similarly digested and ligated into the MBP-TEV-pRSF vector. This plasmid was then transformed into *E. coli* Rosetta 2 cells.

Protein expression and purification. For *Pf*Atg8, *Pf*Atg3, and hLC3, refer to [81]. For hAtg3, refer to Chapter 2.

After transformation of each construct, a single colony was selected, sequence verified, and then grown in Terrific Broth (TB) media with the appropriate antibiotics: 50 μg/mL kanamycin for *Cp*Atg3, *Cp*Atg8, *Et*Atg3, *Et*Atg8, *Nc*Atg3, and *Tg*Atg8; 34 μg/mL chloramphenicol for *Nc*Atg3 and *Tg*Atg8. Growth, induction, harvesting, lysing, and clearance were performed as detailed in Chapter 2.

After lysis and clearance via centrifugation, the supernatant was decanted to incubate with NEB amylose resin for 30 min. Resin was washed with lysis buffer and then incubated for another 30 min with elution buffer: lysis buffer plus 10 mM maltose. The eluted protein was collected and then cleaved overnight with TEV protease at a mass ratio of 1:100, while simultaneously dialyzed into 10 mM HEPES pH 7.5, 100 mM NaCL, 1 mM DTT. Cleaved *Cp*Atg8 was not stable in HEPES buffer, so this dialysis buffer was 10 mM Tris pH 7.5, instead. After overnight cleavage, each protein was separated from MBP and TEV protease by anion exchange using a RESOURCE™ Q column (GE® Healthcare) on an ÄKTA™ purifier system. The Atg8 proteins were collected from the RQ flowthrough, while the Atg3 proteins were collected from a single peak on the chromatogram well-separated from MBP and TEV protease. For SPR, proteins were concentrated to 1-3 mg/mL and buffer exchanged into the indicated running buffer.

Immunoblots. See [83].

SPR PPI assays. An SPR assay was developed for *Pf*Atg8 and *Pf*Atg3 as described previously in [81]. The hAtg3-hLC3 SPR PPI assay setup is described in Chapter 2. Based on these assays, we developed similar interaction assays for *Cp*Atg3-*Cp*Atg8, *Et*Atg3-*Et*Atg8 and *Nc*Atg3-*Nc*/*Tg*Atg8 using our standard running buffer: 10 mM Hepes pH 7.5, 150 mM NaCl, 0.01% Tween 20. All buffers were filtered and degassed for 1 h before each experiment. Small molecule inhibitors were stored in 100% DMSO at 500 μM and tested at 5 μM concentrations (PTA derivatives) or stored at 100 mM and tested at 500 μM concentrations (VLS compounds). The inhibition due to VLS ligands was measured in triplicate. To maintain a constant ratio of DMSO in all small molecule inhibitor assays, the running buffer was supplemented with 1% DMSO. All measurements were analyzed with BioLogic™ Scrubber and corrected for DMSO effects using double referencing with interspersed blank injections and an untreated flow cell on the SPR chip. To control for surface decay, control injections (Atg8 supplemented with DMSO, but no small molecule inhibitors) were injected every 12 injections.

For each Atg3-Atg8 interaction pair, a new CM5 SPR chip was used. For all PPI studies, the Atg3 homologue was immobilized, with the Atg8 partner passed as an analyte over the chip in the presence of small molecule inhibitors or equivalent DMSO. To immobilize Atg3 on the chip, we carried out a pH scouting from pH 4.0 to pH 6.0 in 0.5 pH unit intervals to determine the best conditions for immobilization. Flow cells were activated using N-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) for 7 min with a flow rate of 20 µl/min. For *Cp*Atg3 immobilization, a solution of 50 µg/ml of *Cp*Atg3 in 10 mM sodium acetate pH 5.0 was passed over one flow cell until >500 Response Units (RU) were captured. The same procedure was followed for the immobilization of *Et*Atg3 and *Nc*Atg3, with a pH of 5.0 and pH 5.25, respectively. To block activated free binding sites, we exposed the flow cell to 0.5 M ethanolamine in running buffer for an additional 7 min. For the ALCs, the flow rate was 40 µl/min flow rate with a 75 s contact time, followed by a 1 minute dissociation time and a 25 µl regeneration injection of either 2 M MgCl2 or a mixture of 10 mM NaOH, 10 mM HEPES pH 7.5. For the PTA derivatives screening, the same regeneration conditions were used, but the flow rate was 50 µL/min, and the contact time was 3 min, followed by 30 s wait time and 20 µL regeneration. Dose-response curves were then used to determine the approximate concentration of analyte to obtain a response of 100-150 RU in the absence of ALCs.

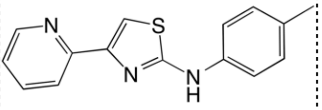
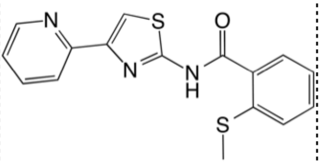
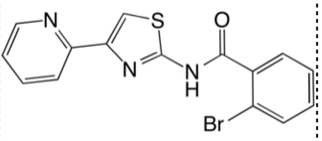
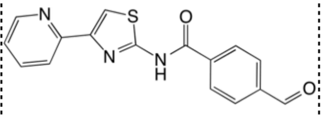
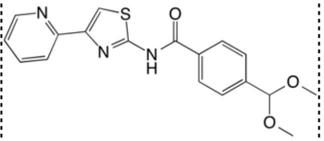
All measurements were carried out on a BiaCore 3000 instrument.

Flow cytometry. We treated uninfected blood or blood infected with *P. falciparum* FCR3-CSA ring-stage parasites. For each sample, 200 µL of culture had drug or 100% DMSO added in 1:100 ratio, for a final concentration of 1% DMSO in all cases. After 16 h incubation, the samples were centrifuged at 500 rcf for 2 minutes, and the supernatant was removed. The pelleted cells were resuspended in 200 µL 1xPBS and then divided into two tubes (one for primary antibody, one for control antibody). The cells were centrifuged again, the supernatant was removed, and they were resuspended in 100 µL of rabbit αVar2csa antibody (1:25 in 1xPBS, 3% BSA) or 100 µL of control rabbit antisera (1:1250 dilution in 1xPBS, 3% BSA) and incubated for 30 minutes at room temperature. The samples were then washed 3x with 600 µL 1xPBS, resuspended in 75 µL Molecular Probes® Alexa Fluor 488 αRabbit IgG (1:500 dilution in 1xPBS, 3% BSA), and incubated at room temperature for 30 minutes. After washing again three times with PBS, the samples were resuspended in 1 mL of DRAQ5 dilution (1:5000 in 1xPBS).

Samples were flown through a BD Biosciences™ FACSCalibur flow cytometer at a rate of 5000 events/s until 10^6 events had been reached for each sample. Data was analyzed using FlowJo™ software (Tree Star, Inc.). Samples were gated for viable cells based on forward and side scatter. Uninfected controls were used to determine the gating for infected (DRAQ5+) cells, and the isotype control was used to determine *Pf*EMP1+/- gating.

Tables

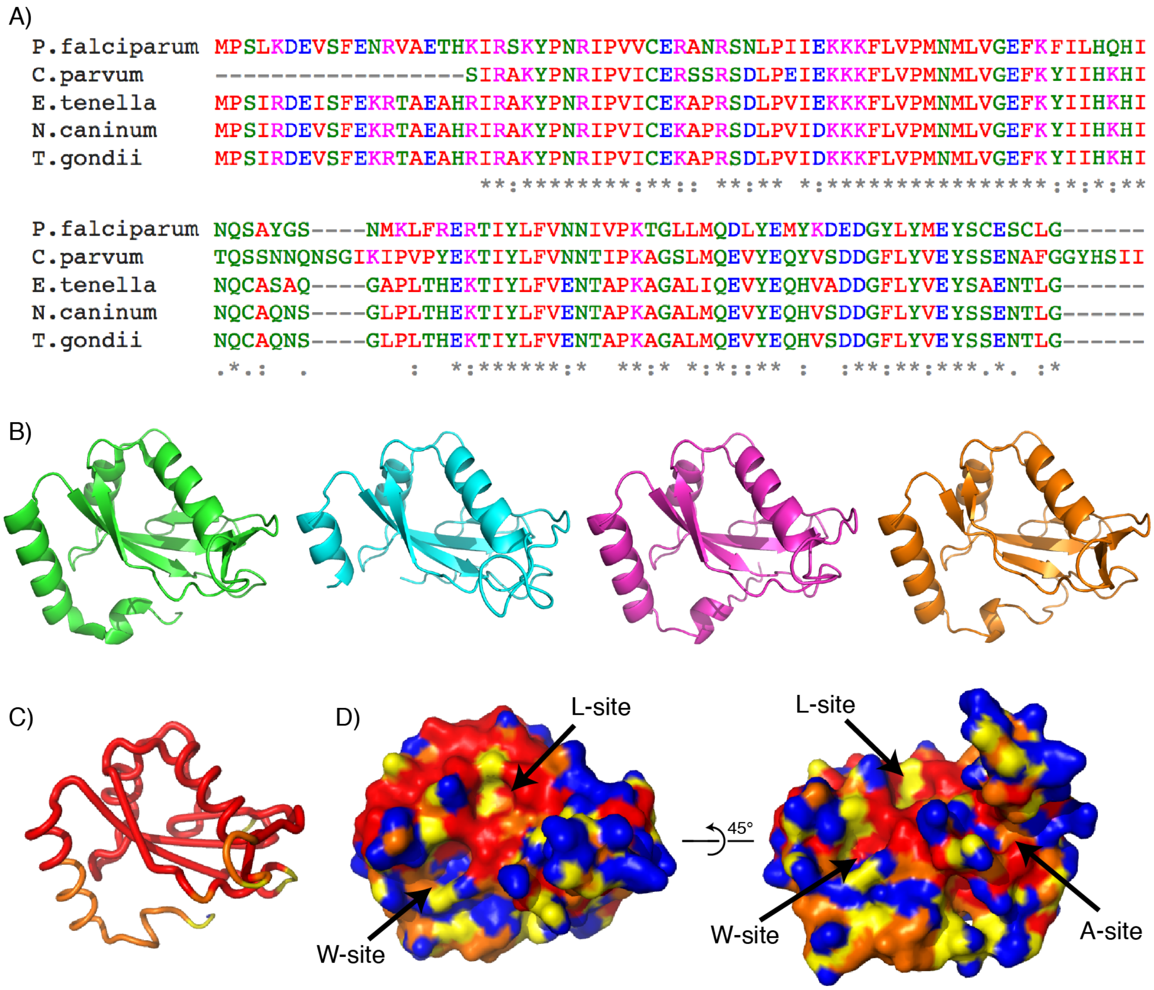
Table 3.1: PTA Compounds and IC_{50} Values

Compound	IUPAC	CID	Structure	Mol Wt	SPR IC_{50} (µM)	Blood stage IC_{50} (µM)
PTA1	N-(4-methylphenyl)-4-pyridin-2-yl-1,3-thiazol-2-amine	746602		267.34	18.16 ± 0.25	1.47 ± 0.31
PTA2	2- methylsulfanyl-N-(4-pyridin-2-yl-1,3-thiazol-2-yl)benzamide	2526359		327.42	14.95 ± 0.22	0.20 ± 0.30
PTA3	2-bromo-N-(4-pyridin-2-yl-1,3-thiazol-2-yl)benzamide)	2454286		360.23	6.03 ± 0.25	1.58 ± 0.35
PTA4	4-formyl-N-(4-pyridin-2-yl-1,3-thiazol-2-yl)-benzamide	N/A		309.34	N/A	N/A
PTA5	4-(Dimethoxymethyl)-N-(4-pyridin-2-yl-1,3-thiazol- 2-yl)benzamide	N/A		355.41	2.86 ± 0.22	1.48 ± 0.23

Figures

Figure 3.1. Apicomplexan Atg8 sequence alignment and homology modeling. (A) Sequence alignment of *Pf*Atg8, *Cp*Atg8, *Et*Atg8, *Nc*Atg8, and *Tg*Atg8 produced by Clustal Omega [125]. Colored by side chain. Symbols correspond to degree of conservation: "*****" indicates identical residues for all homologues; "**:**" indicates high conservation; "**.**" indicates some conservation. Note that *Nc*Atg8 and *Tg*Atg8 are identical. (B) From left to right: the structure of *Pf*Atg8 (in green) followed by the homology models for *Cp*Atg8 (cyan), *Et*Atg8 (magenta), and *Nc/Tg*Atg8 (orange). In (A), it is clear that *Cp*Atg8 is missing 18 residues from the N-terminus as compared to the other homologues, which accounts for the more limited N-terminal helix in the model. (C) The peptide backbone of *Pf*Atg8 colored by conservation via the color_by_conservation.py script [145]. Orientation of the protein is as in (B). Colors range from red (highly conserved) to blue (not conserved) on a rainbow scale and are based on the PyMol alignment RMSD. The backbone is highly conserved among the homologues. (D) Two surface representations of the structure of *Pf*Atg8, including side chains, colored by conservation as in (C). There is a greater degree of variability when considering side chains, but a number of contacts in the W-, L-, and A-sites are still highly or moderately conserved.

Figure 3.1



```
P.falciparum MPSLKDEVSFENRVAETHKIRSKYPNRIPVVCERANRSNLPIIEKKKFLVPMNMLVGEFKFILHQHI
C.parvum     ------------------SIRAKYPNRIPVICERSSRSDLPEIEKKKFLVPMNMLVGEFKYIIHKHI
E.tenella    MPSIRDEISFEKRTAEAHRIRAKYPNRIPVICEKAPRSDLPVIEKKKFLVPMNMLVGEFKYIIHKHI
N.caninum    MPSIRDEVSFEKRTAEAHRIRAKYPNRIPVICEKAPRSDLPVIDKKKFLVPMNMLVGEFKYIIHKHI
T.gondii     MPSIRDEVSFEKRTAEAHRIRAKYPNRIPVICEKAPRSDLPVIDKKKFLVPMNMLVGEFKYIIHKHI
                                **:********:**:: **:** *:****************:*:*:**

P.falciparum NQSAYGS----NMKLFRERTIYLFVNNIVPKTGLLMQDLYEMYKDEDGYLYMEYSCESCLG------
C.parvum     TQSSNNQNSGIKIPVPYEKTIYLFVNNTIPKAGSLMQEVYEQYVSDDGFLYVEYSSENAFGGYHSII
E.tenella    NQCASAQ----GAPLTHEKTIYLFVENTAPKAGALIQEVYEQHVADDGFLYVEYSAENTLG------
N.caninum    NQCAQNS----GLPLTHEKTIYLFVENTAPKAGALMQEVYEQHVSDDGFLYVEYSSENTLG------
T.gondii     NQCAQNS----GLPLTHEKTIYLFVENTAPKAGALMQEVYEQHVSDDGFLYVEYSSENTLG------
             .*.:  .       :  *:******:*  **:* *:*::** :  :**:**:***.*. :*
```

Figure 3.2. Results of VLS compound screen in SPR PPI assay. DMSO control responses are in green. Blue indicates a compound that caused some precipitation of either drug or protein when added to the protein sample. Error bars represent the 95% confidence interval for three measurements. Results are shown for the (A) *C. parvum*, (B) *E. tenella*, and (C) *N. caninum* Atg8-Atg3 interactions. In (B), the SPR responses in the presence of ALC13 and ALC24 are not shown as these compounds produced unreliable SPR responses likely due to precipitation.

Figure 3.2

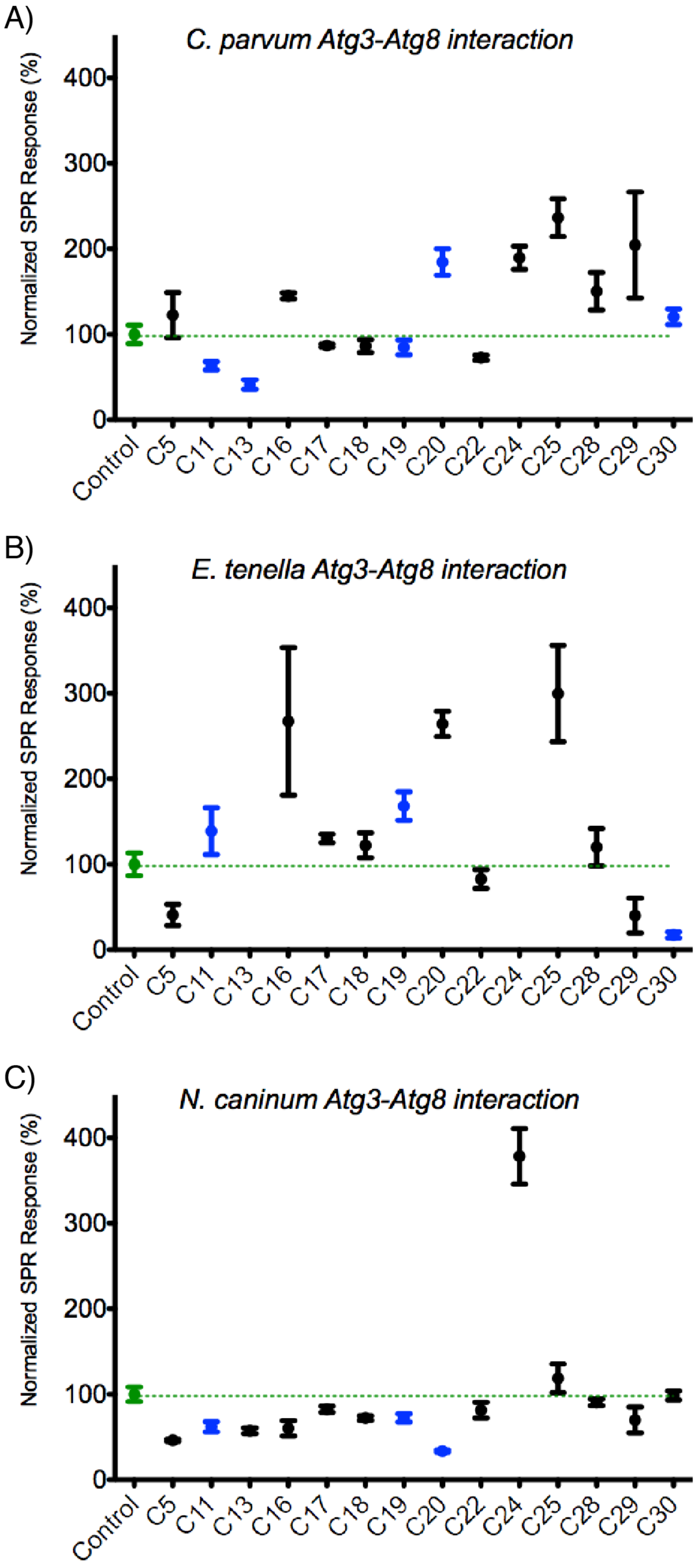

A)
C. parvum Atg3-Atg8 interaction
B)
E. tenella Atg3-Atg8 interaction
C)
N. caninum Atg3-Atg8 interaction
Normalized SPR Response (%)
0
100
200
300
400
Control
C5
C11
C13
C16
C17
C18
C19
C20
C22
C24
C25
C28
C29
C30

Figure 3.3. Western blots demonstrating change in *Pf*Atg8 lipidation state and expression level upon treatment with PTA1. (A) Treatment with varying levels of PTA1 led to a shift from lipidated to unlipidated *Pf*Atg8. (B) Treatment with 50 μM PTA1 for five hours led to a substantial upregulation of *Pf*Atg8 as probed via Western Blot.

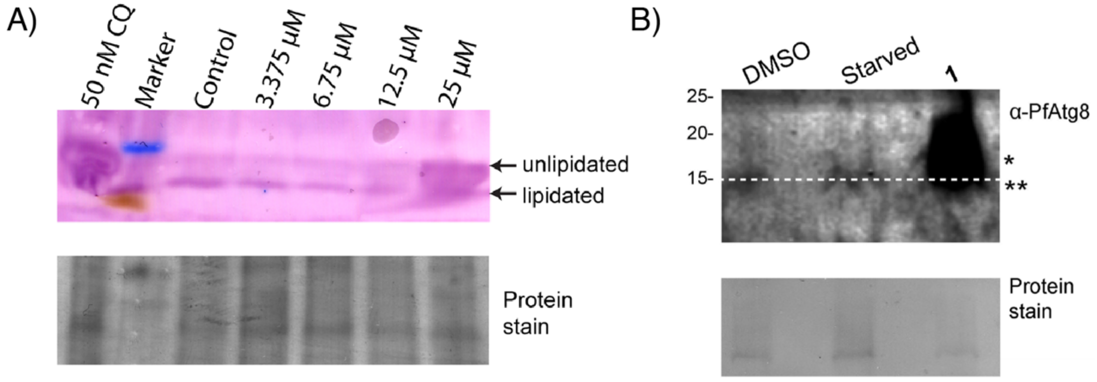


Figure 3.4. PTA-derivative library linkers

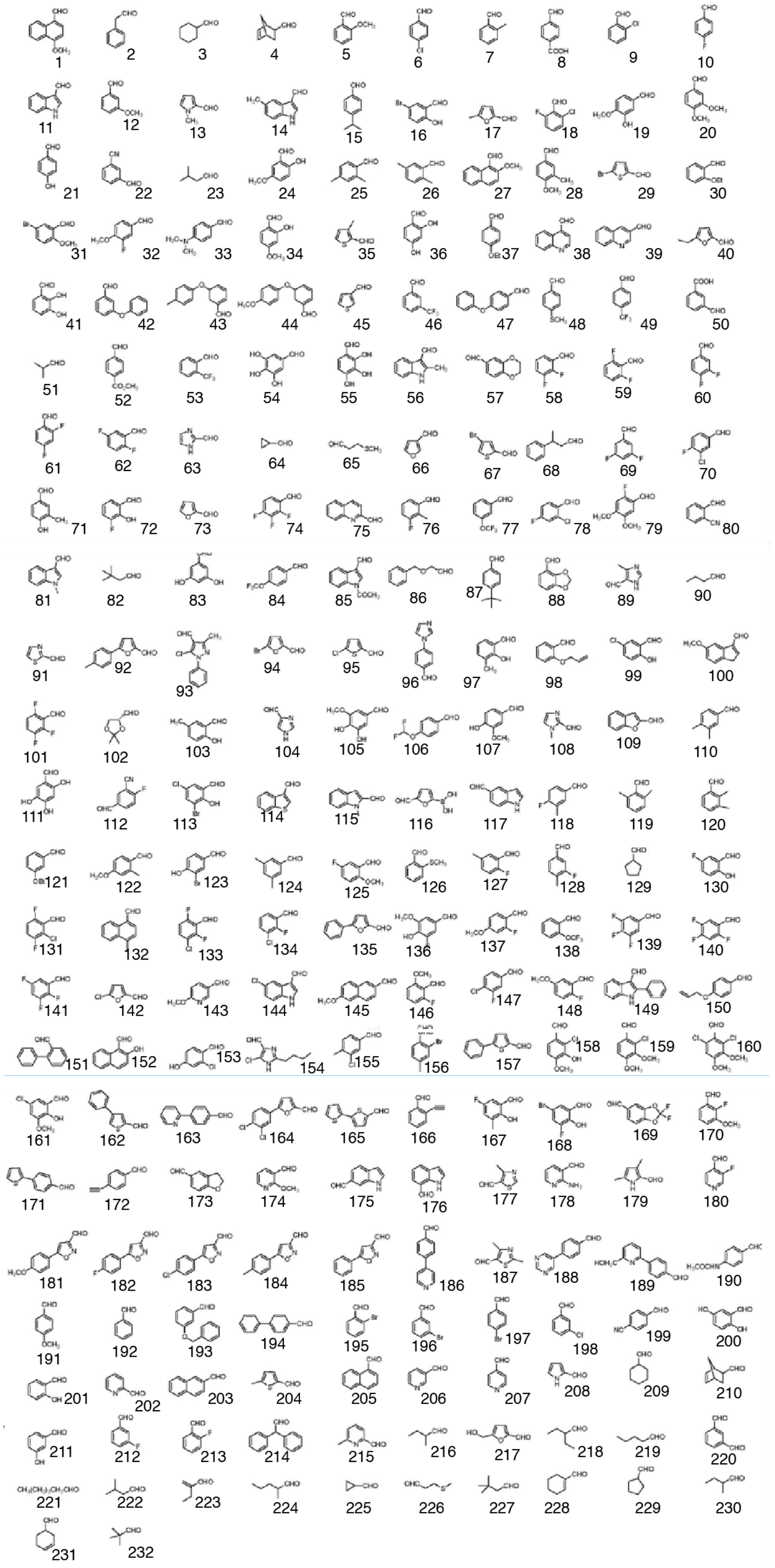


Figure 3.5. Results of PTA derivative library screen in SPR PPI assay. The SPR response in the presence of each analyte is plotted for both human (x-axis) and apicomplexan (y-axis) Atg8-Atg3 interactions. All responses within two standard deviations of the average response are enclosed in a black rectangle. Potential inhibitors are colored red and enclosed in the green rectangle, representing analytes that don't significantly affect the the human interaction but produce a significant reduction in the SPR response for the apicomplexan interaction. The results are shown for (A) the *P. falciparum* Atg8-Atg3 interaction with the PTA derivatives tested at 0.5 μM and (B) the *C. parvum*, (C) *E. tenella*, and (D) *N. caninum* Atg8-Atg3 interactions with the derivative library tested at 5 μM.

Figure 3.5

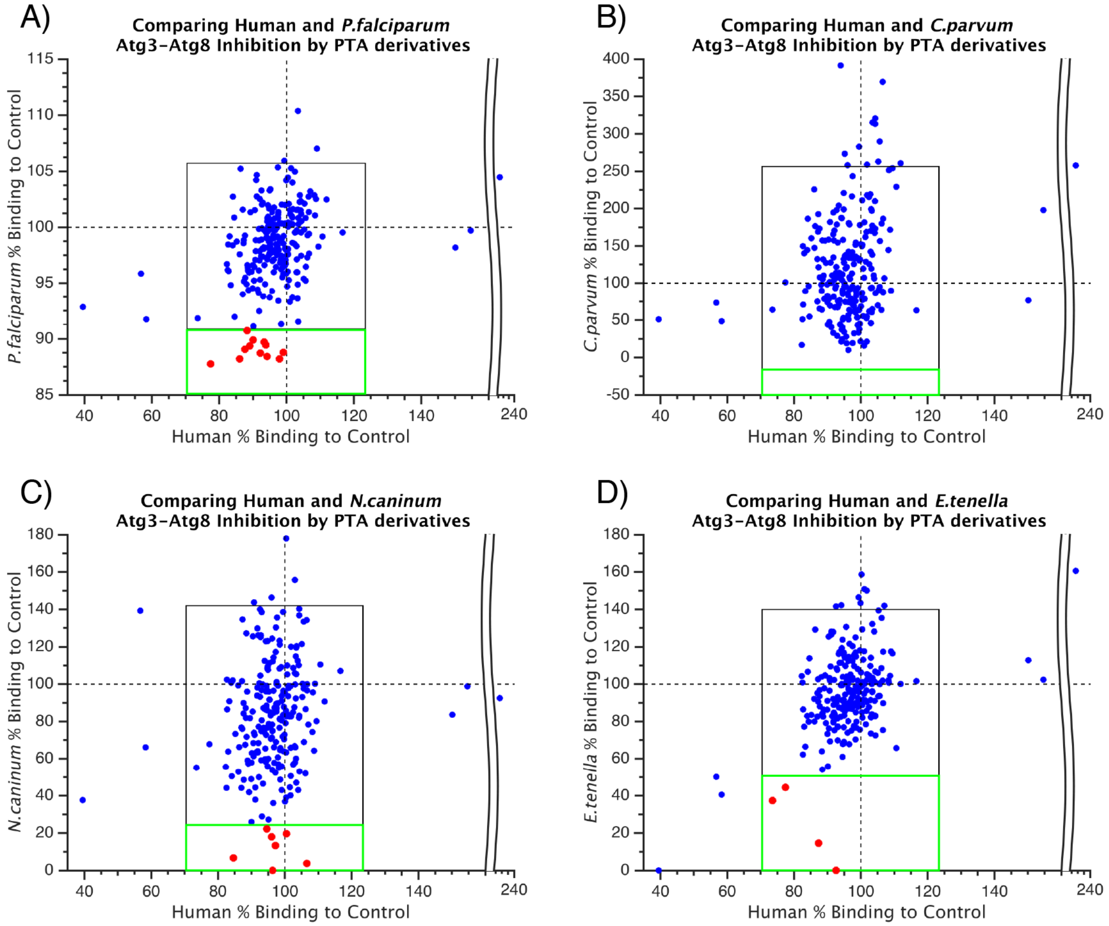
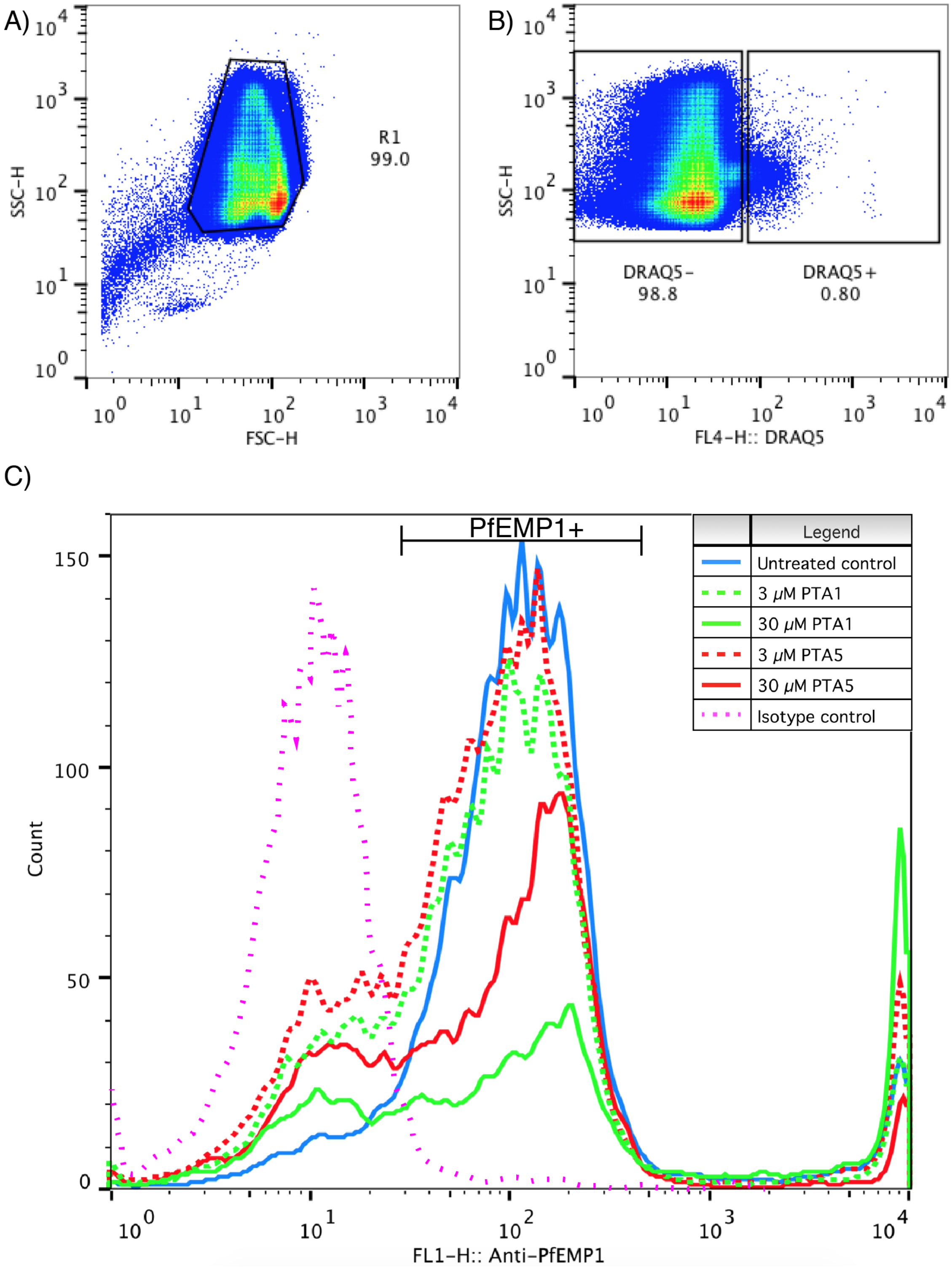

A)
Comparing Human and P.falciparum
Atg3-Atg8 Inhibition by PTA derivatives
P.falciparum % Binding to Control
115
110
105
100
95
90
85
40
60
80
100
120
140
240
Human % Binding to Control
B)
Comparing Human and C.parvum
Atg3-Atg8 Inhibition by PTA derivatives
C.parvum % Binding to Control
400
350
300
250
200
150
100
50
0
-50
40
60
80
100
120
140
240
Human % Binding to Control
C)
Comparing Human and N.caninum
Atg3-Atg8 Inhibition by PTA derivatives
N.caninum % Binding to Control
180
160
140
120
100
80
60
40
20
0
40
60
80
100
120
140
240
Human % Binding to Control
D)
Comparing Human and E.tenella
Atg3-Atg8 Inhibition by PTA derivatives
E.tenella % Binding to Control
180
160
140
120
100
80
60
40
20
0
40
60
80
100
120
140
240
Human % Binding to Control

Figure 3.6. Flow cytometry analysis of *Pf*EMP1 levels in treated vs untreated iRBCs. (A) An example dot plot illustrating the gating used to select viable single cells based on forward (x-axis) and side (y-axis) scatter. (B) An example dot plot illustrating the gating for DRAQ5+ (infected) cells. (C) The isotype control (magenta) and the infected, untreated control were used to determine *Pf*EMP1+/- gating. For viable, infected cells, levels of *Pf*EMP1 differed between the cells that were left untreated and cells that were treated with 3 or 30 μM PTA1 or PTA5.

Figure 3.6.


A)
SSC-H
FSC-H
R1
99.0
B)
SSC-H
FL4-H:: DRAQ5
DRAQ5-
98.8
DRAQ5+
0.80
C)
PfEMP1+
Count
FL1-H:: Anti-PfEMP1
Legend
Untreated control
3 μM PTA1
30 μM PTA1
3 μM PTA5
30 μM PTA5
Isotype control
150
100
50
0
10^0
10^1
10^2
10^3
10^4

CHAPTER 4

Conclusions and Future Directions

Summary

The parasitic protists of the phylum, Apicomplexa, represent an area of active research. Chapter 1 gives a broad overview of this phylum and what is currently known about apicomplexan autophagy. From the research that has been conducted so far, it appears that these parasites have co-opted the autophagy pathway to serve novel purposes unique to the parasite lifecycle; for example, Atg8 localizes to the apicoplast, an organelle that does not exist in higher level eukaryotes. Also in this chapter, the opportunity to target the Atg3-Atg8 protein-protein interaction (PPI) within apicomplexans in general and within *Plasmodium* specifically is discussed in detail. In both *P. falciparum* and *T. gondii*, Atg8 has been found to be essential. In addition, the structure of *Pf*Atg8 is known and reveals an apicomplexan loop (A-loop) that represents a point of divergence from human Atg8 homologues but is conserved in Apicomplexa.

Chapter 2 presents the results of targeting this A-loop in a recent virtual ligand screening (VLS) conducted against the A-loop pocket in *Pf*Atg8. From this screen, apicomplexan loop compound 25 (ALC25) showed dose dependent inhibition of the *Pf*Atg3-*Pf*Atg8 PPI in SPR with an IC_{50} of 18.5 μM and had a similar *in vivo* IC_{50} of 19 μM in a SYBR green I assay. ALC25 inhibition also produced a shift in *Pf*Atg8 lipidation and an upregulation of *Pf*Atg8 production, as had previously been shown for another *Pf*Atg3-*Pf*Atg8 PPI inhibitor [83]. Our follow-up SPR and ITC experiments confirmed that ALC25 does not bind to the W- or L-sites and furthermore does not bind to the loop deletion mutant of *Pf*Atg8. These results taken together point to ALC25 inhibiting parasite growth by binding to the A-loop pocket of *Pf*Atg8, disrupting the Atg3-Atg8 PPI, proving that this is a viable approach to drug development against *P. falciparum*.

Finally, in Chapter 3, I present the unpublished results of our efforts to target apicomplexan Atg3-Atg8 interactions with two classes of ligands targeting the *Pf*Atg3-*Pf*Atg8 interaction. From these screens, we identify a number of potential hits inhibiting each of the homologous Atg3-Atg8 interactions: *P. falciparum*, *C. parvum*, *E. tenella*, *N. caninum*/*T. gondii*. In this chapter, also discussed are the exciting results of our *Pf*EMP1 flow cytometry assay, in which inhibitors of the *Pf*Atg8-*Pf*Atg3 PPI lead to an overall decrease in the level of surface-exposed *Pf*EMP1, a major virulence factor.

Future directions and remaining questions

While a number of studies have been conducted to determine the roles of *Tg*Atg8 and *Pf*Atg8, many questions remain. Recent studies have explored the functions of these proteins in the parasite lifecycle, finding evidence for a wide variety of roles (see Chapter 1). Though the apicomplexan life cycle and Atg8 sequence similarities suggest that this protein may also play a significant role in other apicomplexans, its function has not yet been investigated outside of *Plasmodium* and *Toxoplasma*. Verification of the hits pursued in Chapter 3, which show preliminary inhibition of the Atg3-Atg8 PPI in *C. parvum*, *E. tenella*, and *N. caninum*, may provide useful tools for investigating the importance of the Atg8 conjugation pathway in some of these organisms. If Atg8 essentiality holds in these species as well, there is great potential for a novel class of anti-apicomplexan therapeutics among the PTA-derivative molecules.

Of great significance as well are the results of the *Pf*EMP1 flow cytometry experiment implying a role for *Pf*Atg8 in export to the host cell. As discussed in Chapter 3, however, more work needs to be done to explore this connection. Better co-localization immunofluorescence assays may be possible with higher resolution techniques and more specific antibodies, as we now have an antibody directly targeted against *Pf*Atg8, versus the less specific *Tg*Atg8 antibody that had previously been used. Additionally, flow cytometry and western blot analysis of global expression levels of PfEMP1 would answer important questions about the mode by which PTA1 and PTA5 reduce surface-level expression of the protein. If *Pf*Atg8 is, indeed, involved in export of *Pf*EMP1 to the host, then inhibitors of the Atg3-Atg8 PPI would have important implications for severe, cerebral malaria.

Finally, based on the work presented here, an important potential area of lead development is the linking of the two classes of molecules explored in this thesis. Our data support the hypothesis that ALC25 is indeed binding to the A-loop pocket as its primary mode of inhibition, while the PTA derivatives are thought to bind to the W- and L-sites. To take advantage of the tighter binding of the PTA derivatives, but improve specificity for the plasmodial target, it may be possible to connect the molecules with a variety of linkers, allowing for a drug that took advantage of both sites. To probe the feasibility of such an endeavor, we could conduct SPR PPI assays using both drugs to determine if they inhibit synergistically or antagonistically. In general, drug development efforts would be enhanced by the solution of ligand-bound crystal structures.

Developing cost-effective anti-apicomplexan therapeutics

Recent years have seen an increase in malaria funding globally and a corresponding decline in rates of infection and mortality [146]. Infection prevalence in endemic Africa, in particular, has halved over the last decade [3], giving great hope for the possibility of malaria eradication. While this trend is encouraging, we must be realistic in facing remaining challenges. Malaria is still a serious threat, and many figures regarding global malaria mortality may be underestimated: by some measures, the current rates of death as a result of malaria infection are still likely to be above one million per year [147]. Additionally, in comparison to its impact on global health, malaria is still underfunded, especially when compared to diseases such as cancer, for example.

Furthermore, as discussed in previous chapters, the supply of effective antimalarials is steadily dwindling. For many other apicomplexan diseases, the arsenal of available therapeutics is even more limited, nonspecific, or toxic [33,148]. Such realities highlight the need for new effective medication, both in treating malaria and in confronting other apicomplexan ailments.

Unfortunately, for diseases like malaria that predominantly affect citizens of developing nations, there is little incentive for pharmaceutical companies to invest in research and development when there is no assurance of long-term profits. The pharmaceutical industry often takes a high-throughput screening approach to drug development, synthesizing and testing impressively large libraries of potential drugs [149]. These screens are conducted in the hopes of discovering a hit that can then be marketed at a price to recuperate the enormous expense of the discovery process, now estimated to be over US$2 billion [150]. Such an approach relies on the profits from successful drugs to compensate for the expense of failed leads [150,151]. Inherent to this equation is the assumption of a wealthy target population. However, for diseases like malaria, there is no guarantee that those who suffer from it will be able to foot the bill.

Because of this challenge and the general difficulty in producing new drugs [152], there is significant appeal for the methods featured in this thesis – namely, VLS, the MMV Malaria Box, and screening against the Atg3-Atg8 PPI in multiple organisms with related compounds. In certain situations, VLS alone has the potential to reduce drug development costs by up to 50% in the long run [153]. Similarly, the Malaria Box can substantially lower long-term costs by providing researchers with a known starting place: each of the compounds in the screen has confirmed activity against blood-stage *P. falciparum* and is commercially available [154].

Finally, the ability to target multiple Atg3-Atg8 PPIs with one class of molecules is desirable on multiple levels. Recent years have seen an increasing awareness of the value of drug repositioning [155]. Wherever it is possible, drug repositioning represents a significantly cheaper alternative to the traditional drug discovery process. Along the same lines, drug development of a class of molecules targeting multiple apicomplexan PPIs would limit the inherent expense of the drug development process. It could also allow for the potential treatment of multiple zoonotic infections simultaneously, an approach that is already being pursued by others in the field [156].

APPENDIX A

Crystallizing MBP

Introduction

The field of crystallography has seen many (now notorious) examples of flawed, wildly inaccurate, or completely fake crystal structures (see [157] for early examples, [158] and [159] for more recent examples). Luckily, with the advent and increased utility of numerous crystallographic validation tools, the modern crystallographer is more or less saved from such a fate. High free R factors, poor geometry, anomalous Ramachandran plots, numerous steric clashes, and bad rotamers, among other features, sound the alert as to possible problems. Many of these validation checks are automated into the refinement process, allowing even beginners to reliably check the quality of their structure, though the odd persevering researchers may still find their way down the rabbit hole, as in [158]. The following description of a failed crystallographic effort is in a way a testament to this, as the false structure never made it past the refinement stage due to abnormal crystal contacts, poor geometry and a consistently high free R-factor. Nevertheless, many of the other best practices of crystallography were not followed, practices that might have prevented many wasted hours spent in fruitless attempts at structure solution. Hence, for the purposes of self-edification and as a cautionary tale for other graduate students, the details of this effort and possible precautions that could have been followed have been included herein.

Materials, Methods, & Possible Precautions

Strains & Growing Conditions. The details for the constructs and growing conditions for *Cp*Atg8, *Et*Atg8, and *Tg*Atg8 are listed in Chapter 3. In summary, all three proteins were encoded on a modified pRSF vector with an N-terminal Maltose Binding Protein (MBP) followed by a Tobacco *Et*ch Virus (TEV) cleavage site. *Cp*Atg8 and *Et*Atg8 were expressed in BL21 *E. coli* cells and *Tg*Atg8 was expressed in Rosetta 2 cells. They were grown in Terrific Broth (TB) at 37°C to an OD_{600} of 3 and then induced with 0.5 mM IPTG and grown overnight at 20°C.

Purification of MBP fusion proteins. After harvesting, lysis, and clearance (see chapters 2 and 3), the cleared lysate was incubated with New England Biolabs® amylose resin for 30 minutes. The flowthrough was collected (flowthrough – Figure A.1A,#1) and then the beads were washed with lysis buffer (10 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT) (wash – Figure A.1A,#2). Bound Atg8 protein was incubated with lysis buffer plus 10 mM maltose for 30 minutes and then collected (elution – Figure A.1A,#3-6). This eluted protein was pooled and cleaved overnight at 4°C with TEV protease at a 1:100 mass ratio (cleaved protein – Figure A.1B,#1) and dialyzed into Buffer A: 10 mM HEPES pH 7.5 (Tris pH 7.5 for *Cp*Atg8), 75 mM NaCl, 1 mM DTT.

The Atg8 protein was separated from MBP and TEV protease via anion exchange using a GE® Healthcare RESOURCE™ Q column, using a gradient to elute various proteins going from the low salt Buffer A condition to the high salt Buffer B: 10 mM HEPES pH 7.5 (Tris pH 7.5 for *Cp*Atg8), 1 M NaCl, 1 mM DTT. The Atg8 proteins have high isoelectric points, around or above pH 8, so they were not negatively charged at a pH of 7.5. This means that they did not bind to the Q column, and were instead collected from the flow-through coming through the column (Figure A.1C,D#1-3). MBP and TEV protease did predominantly bind to the column and were eluted in later peaks (Figure A.1C,D#5-10).

Precautions. While Figure A.1 suggests that the Atg8 protein was reasonably pure before being set in crystal trays, the purity of the Atg8 proteins varied between preps. A solution to this impurity, and specifically contamination with MBP, would be to add an additional purification step. Early efforts to use size exclusion chromatography to this effect were unsuccessful, as it consistently co-eluted with MBP, likely because of Atg8 dimerization. However, there are other ways to improve purity, including hydrophobic interaction chromatography, repeated ion exchange chromatography, additional affinity tag chromatography (through the addition of an N-terminal His-tag for example), along with many others. A simple starting point can be found in [160], which provides a number of successful strategies used by structural genomics centers to crystallize proteins. For a simple listing of possible purifications strategies, consider the GE® Healthcare Strategies for Protein Purification Handbook or, for more depth, consider a textbook such as [161].

Setting crystal trays. The protein collected from the flowthrough of the anion exchange was concentrated to approximately 5 mg/mL, until protein aggregate was noted. The concentrated protein was then centrifuged at 18000 *g* to remove protein aggregates. These proteins were then set in 3x96-well sitting drops using a variety of crystallographic screens at a 2:1 ratio, 400 nL of protein and 200 nL of crystal reservoir. The top drop had *Cp*Atg8, the middle drop had *Et*Atg8, and the bottom drop had *Tg*Atg8.

Crystals. From the initial screens, crystal formation was noted in the Jena Bioscience JBScreen Classic blocks from the reservoir condition that contained 100 mM Tris pH 8.5, 0.2 M $MgCl_2$, 30% w/v PEG-4000 (Figure A.2A). This crystallization was reproduced at pH values varying from 7 to 8, PEG-4000 or PEG-3350 concentrations varying from 25% to 35%, and with various additives in 4x24 sitting well formats. All of the crystals were plate-like and had some mosaicity, but were untwinned.

Precautions. Very simply, the crystals could have been run on SDS-PAGE (even after being shot) and thus discriminated based on the protein bands that appeared on the gel. For a simple description, see *e.g.*, http://proteincrystallography.org/ccp4bb/message35357.html

Data collection, indexing and processing. The crystals from the initial screen were shot at Brookhaven National Laboratories, using iMosflm for the data collection strategy. The diffraction data were indexed and processed using XDS [162] via XDSGUI. The crystals were identified via XDS and Pointless to be in space group 19, $P2_12_12_1$, orthorhombic with unit cell dimensions approximately: 48.72 Å×57.66 Å×123.20 Å. Some of the crystals out of the initial screen had sub 2 Å resolution according to the XSCALE data summary (Figure A.2B). Analyzing the Matthew's coefficient for one of the "*Et*Atg8" crystals with CCP4's Matthew_coef yielded: for two *Et*A*Tg*8 molecules in the asymmetric unit (ASU), a coefficient of 3.02, solvent content of 59.3%, and probability of 40%; for three molecules in the ASU, a coefficient of 2.02, solvent content of 39.0%, and probability of 59%.

Precautions. The Matthew's coefficient for MBP was 1.99 for one molecule in the ASU, with a solvent content of 38.36%, and a probability of 100% as no other configurations were possible given the unit cell dimensions. Thus, an analysis of the Matthew's coefficient alone was not enough to decisively rule out either the Atg8 homologue or MBP.

Molecular Replacement. After data processing, all efforts to use molecular replacement failed or produced dubious results. All three homologues (*Cp*Atg8, *Et*Atg8, and *Nc*/*Tg*Atg8) have greater than 60% sequence identity with *Pf*Atg8 (Figure 3.1A), making any crystal structures of these proteins good candidates for molecular replacement with the known structure of *Pf*Atg8. However, all attempts at molecular replacement with Phenix Phaser-MR simply never finished, as it continued to search for solutions meeting certain criteria and failed to find any. CCP4's MolRep produced solutions, but with a Contrast of less than 3, a wRfac of 1 or greater, and a Score of 0.1-0.3, all of which point to the solution being false. Molecular replacement was also attempted using the same homology models as in Chapter 3, but similarly failed or produced weak results.

Precautions. Once molecular replacement had failed with the *Pf*Atg8 structure and the other homology-modeled structures, the next logical step would have been to attempt molecular replacement with likely contaminants such as MBP and TEV protease. In fact, this is how the mistake was eventually discovered.

Refinement. Some of the more plausible solutions from the *Et*Atg8 molecular replacement via MolRep were run through rigid body refinement and real space refinement using CCP4's Refmac5 and further refinement with phenix.refine. But inevitably, the R_{free} value would get stuck. In the case where refinement was taken the furthest, the R_{work} got down to 0.25, but the free R factor was stuck above 0.38. In every case, visual inspection of the solution would reveal unnatural crystal packing: either the molecules overlapped or there were large empty planes devoid of molecules. Additionally, the validation tools available through Coot and Phenix consistently produced uninspiring results.

Precautions. By its very nature, the free R factor (R_{free}) serves as a precaution against finding your signal in the noise. To give a brief overview, the R factor is defined as $R = \frac{\sum|F_{obs}-F_{calc}|}{\sum|F_{obs}|}$, where the F_{obs} values are the observed structure factor amplitudes and the F_{calc}'s are the structure factor amplitudes derived from the model. Essentially, the R factor is a measure of agreement between the observed data and the data calculated based on the model.

However, in protein crystal structure refinement, there are many more fitting parameters than observations, leading to the possibility of overfitting, in which your model does not represent the structure that generated the data but nevertheless accurately reproduces the observations. A simple way to detect overfitting is to check the predictive power of your model: leave out 5% of your reflections when generating your model and then check to see if your model accurately predicts this excluded data. This is precisely what the R_{free} is. The equation is the same as for the R factor, but the F_{obs} values are structure factor amplitudes for a random subset of reflections that have been left out of the refinement process. When your model is gradually tweaked to fit the data, instead of generated from the data, it will result in a high R_{free}: your model will fail to be predictive of the reflections that were left out of the refinement process. This value is compared to the R factor for the reflections that were used in calculating the model, called the R_{work}.

In this case, the model did not accurately describe the data. While other problems with the diffraction data can lead to a problem with the R_{free}, a high R_{free} is a classic symptom of an incorrect model. The high R_{free} combined with clearly incorrect crystal contacts points toward obvious problems with the model.

Further efforts. Because of the high resolution and the possible solutions, I continued to try to refine the model, assuming the problems arose from undetected twinning or other impenetrable molecular replacement issues. I was encouraged in these efforts by some favorable preliminary electron density maps (Figure A.2C).

MBP solution. After a certain amount of time, however, it became clear that nothing was working. As a result, I backtracked through my efforts and attempted molecular replacement with MBP, at which point the actual solution became easily apparent. Molecular replacement results were definitive and after only two rounds of refinement through Phenix.refine, the R_{work} and R_{free} values were 21% and 25%, respectively.

Conclusion

Crystallography can be a taxing and time-consuming endeavor. This process can be made easier by observing best practices or more difficult by circumventing these checks in an effort to save time. In the end, however, performing the necessary tests to ensure the reality of what you are investigating can save much wasted time. I hope the above guide provides examples of how to avoid the problems that I faced.

Figures

Figure A.1. Example *Cp*Atg8 purification. (A) A gel showing the flowthrough from binding to amylose resin (lane 1), the buffer wash of the amylose resin (lane 2), and the maltose elutions (lanes 3-6) from a *Cp*Atg8 purification. The ladder in lane 7 is the Biorad Color Plus Prestained Protein Ladder. (B) Lane 1 shows the cleavage efficiency after overnight TEV protease cleavage of the MBP-TEV-*Cp*Atg8 construct. Lane 2 – ladder. (C) The chromatogram from the anion exchange. The blue curve (280 nm absorbance) shows where protein is eluted from the column; the red curve (260 nm absorbance) shows a peak in DNA elution in fractions D5-E2; green is the concentration of the high salt buffer B (from 0-100%); brown shows the conductivity – the actual concentration of salt lags behind the green curve. Stars correspond to fractions run on the gel in (D). (D) Lanes 1-3 correspond to the flowthrough, containing predominantly cleaved *Cp*Atg8 – fractions B12, B6, and B1. Lane 4 is the ladder. Lanes 5-10 correspond to the protein peaks that elute off of the column with higher salt – fractions C7, C9, C11, D12, D10, D8.

Figure A.1

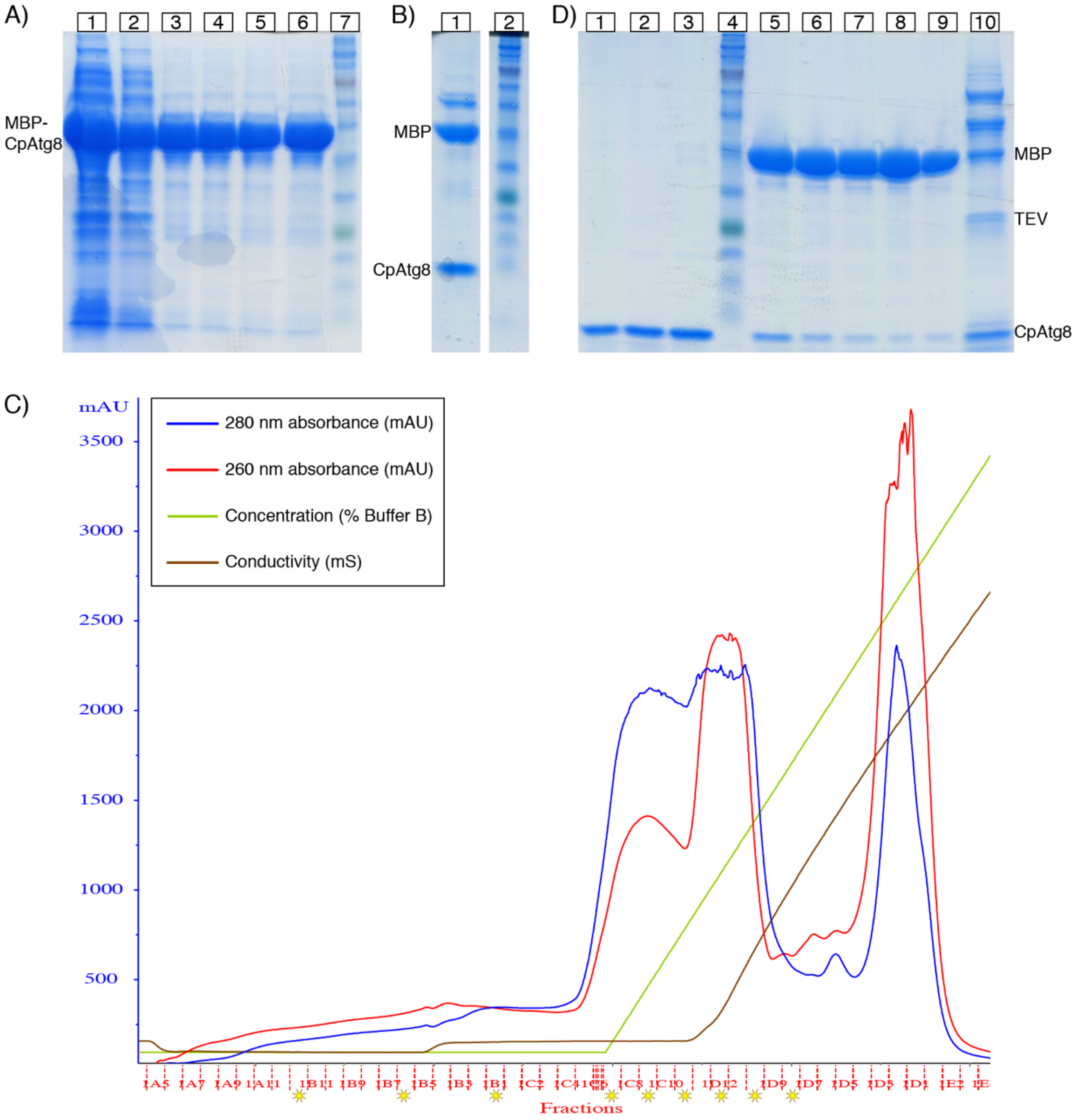
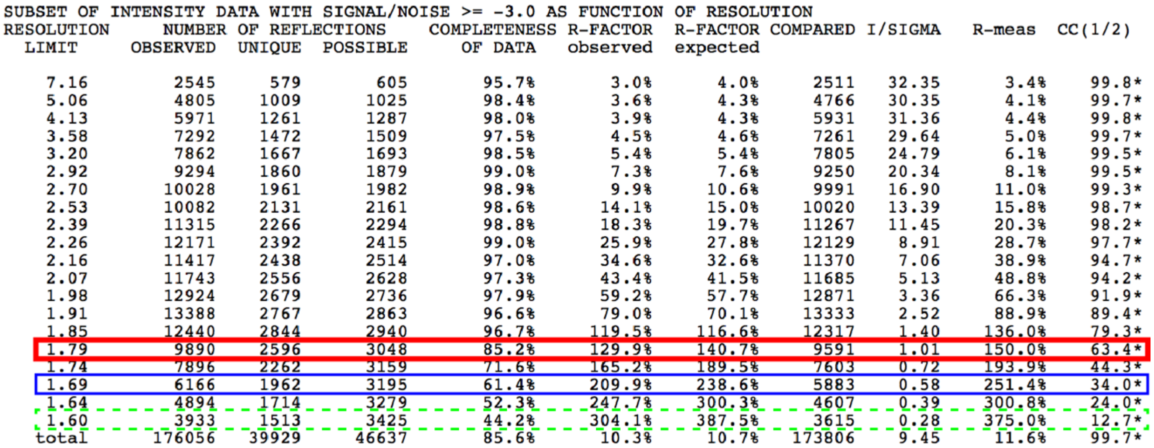

A)
1
2
3
4
5
6
7
MBP-
CpAtg8
B)
1
2
MBP
CpAtg8
D)
1
2
3
4
5
6
7
8
9
10
MBP
TEV
CpAtg8
C)
mAU
3500
3000
2500
2000
1500
1000
500
280 nm absorbance (mAU)
260 nm absorbance (mAU)
Concentration (% Buffer B)
Conductivity (mS)
1A5
1A7
1A9
1A11
1B11
1B9
1B7
1B5
1B3
1B1
1C2
1C4
1C6
1C8
1C10
1D12
1D9
1D7
1D5
1D3
1D1
1E2
1E
Fractions

Figure A.2. MBP crystal details. (A) Plate-like clusters from initial screen. (B) XSCALE summary of crystal statistics in each resolution range. Based on CC1/2 values, there is significant data down to a resolution of 1.6 Å; based on an I/sigma cutoff of 0.5, there is significant data down to 1.69 Å. However, the completeness is limited in these resolution shells, and becomes more reliable at around 1.8 Å. (C) Initial misleading electron density for the false *Et*Atg8 molecular replacement model.

Figure A.2

A)

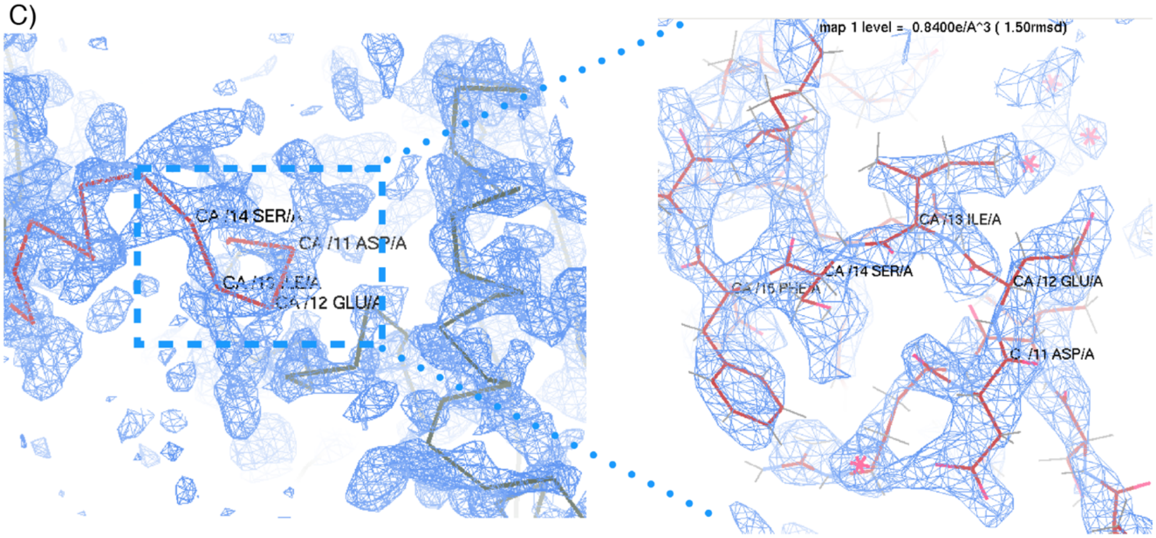
B)



```
SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION
RESOLUTION     NUMBER OF REFLECTIONS    COMPLETENESS R-FACTOR  R-FACTOR COMPARED I/SIGMA   R-meas  CC(1/2)
  LIMIT     OBSERVED  UNIQUE  POSSIBLE     OF DATA   observed  expected

    7.16        2545     579       605       95.7%       3.0%      4.0%      2511   32.35      3.4%    99.8*
    5.06        4805    1009      1025       98.4%       3.6%      4.3%      4766   30.35      4.1%    99.7*
    4.13        5971    1261      1287       98.0%       3.9%      4.3%      5931   31.36      4.4%    99.8*
    3.58        7292    1472      1509       97.5%       4.5%      4.6%      7261   29.64      5.0%    99.7*
    3.20        7862    1667      1693       98.5%       5.4%      5.4%      7805   24.79      6.1%    99.5*
    2.92        9294    1860      1879       99.0%       7.3%      7.6%      9250   20.34      8.1%    99.5*
    2.70       10028    1961      1982       98.9%       9.9%     10.6%      9991   16.90     11.0%    99.3*
    2.53       10082    2131      2161       98.6%      14.1%     15.0%     10020   13.39     15.8%    98.7*
    2.39       11315    2266      2294       98.8%      18.3%     19.7%     11267   11.45     20.3%    98.2*
    2.26       12171    2392      2415       99.0%      25.9%     27.8%     12129    8.91     28.7%    97.7*
    2.16       11417    2438      2514       97.0%      34.6%     32.6%     11370    7.06     38.9%    94.7*
    2.07       11743    2556      2628       97.3%      43.4%     41.5%     11685    5.13     48.8%    94.2*
    1.98       12924    2679      2736       97.9%      59.2%     57.7%     12871    3.36     66.3%    91.9*
    1.91       13388    2767      2863       96.6%      79.0%     70.1%     13333    2.52     88.9%    89.4*
    1.85       12440    2844      2940       96.7%     119.5%    116.6%     12317    1.40    136.0%    79.3*
    1.79        9890    2596      3048       85.2%     129.9%    140.7%      9591    1.01    150.0%    63.4*
    1.74        7896    2262      3159       71.6%     165.2%    189.5%      7603    0.72    193.9%    44.3*
    1.69        6166    1962      3195       61.4%     209.9%    238.6%      5883    0.58    251.4%    34.0*
    1.64        4894    1714      3279       52.3%     247.7%    300.3%      4607    0.39    300.8%    24.0*
    1.60        3933    1513      3425       44.2%     304.1%    387.5%      3615    0.28    375.0%    12.7*
   total      176056   39929     46637       85.6%      10.3%     10.7%    173806    9.45     11.6%    99.7*
```

C)



APPENDIX B

Virtual Ligand Screening against Eukaryotic Topoisomerase IIA

Introduction

DNA topoisomerases solve important topological problems resulting from a variety of DNA transactions [163,164]. Type IIA topoisomerases (TOP2s), in particular, regulate DNA topology by passing one segment (the T-segment) of DNA through a transient double strand break in another segment (the G-segment) (Figure B1.A) [165,166]. This action allows the enzyme to relax superhelical tension and disentangle chromosomal segments [167]. The enzyme, which is a homodimer in eukaryotes, accomplishes this complex operation through the coordinated motion of dimerizing subunits [168–170]. The dimer interfaces make up three “gates:” the N-gate, the DNA gate, and the C-gate (Figure B.1A,B) [163].

Because of the essential role TOP2s play in DNA maintenance and cell division, inhibitors of these enzymes are important antibacterial and anticancer agents [171,172]. ICRF-187, also known as dexrazoxane, is a known inhibitor of eukaryotic TOP2s [173,174]. In particular, its ability to bind the alpha isoform in humans (hTOP2A) makes it relevant to cancer treatment. There are two TOP2 isoforms in humans – alpha and beta (hTOP2B) – which have differential expression throughout the body, with hTOP2A being overexpressed in tumor cells [175,176].

A crystal structure of dexrazoxane bound to *Saccharomyces cerevisiae* TOP2 (ScTOP2) was solved to 1.9Å resolution, revealing its mode of inhibition [177]. Dexrazoxane binds to the interface between the two lobes of the ATPase domain, the N-gate, trapping the molecule in a closed state and preventing the overturn of ATP molecules necessary for the function of the protein (Figure B.1B).

While dexrazoxane is a useful tool in probing TOP2 function, it has limited efficacy as an anti-cancer drug due to its high (µM) *in vivo* IC_{50} [178]. Additionally, it is not known whether dexrazoxane preferentially binds hTOP2A over hTOP2B. Inhibition of hTOP2B can result in off-target effects, including cardiotoxicity and secondary malignancies [171,179,180].

Because there is room for improvement, this represents a case for which virtual ligand screening (VLS) is an attractive option; specifically, the target structure is known and has been solved with a bound ligand. This provides a starting location for virtual docking and an initial structure for use in enriching a docking library. To maximize the likelihood of identifying positive hits, three diverse docking programs were used in an initial docking to identify the best program for this target: Schrödinger® Glide, OpenEye™ Fred, and AutoDock™ Vina.

The following work details the efforts to identify novel ligands inhibiting the eukaryotic TOP2 ATPase domain through VLS and *in vitro* verification.

Results

1QZR selected as target structure, prepared for docking

There are several structures of eukaryotic TOP2 in the protein data bank (PDB), but only five of the ATPase domain. These include two crystal structures from yeast (PDB ID: 1PVG and 1QZR) and three from human (PDB ID: 1ZXM, 1ZXN and 4R1F). 1QZR, solved to 1.9 Å, features the ATPase domain of ScTOP2 bound to dexrazoxane, and is the only structure crystallized with a bound inhibitor. 1PVG is the apo structure of the same construct.

A comparison of 1QZR to the human structure 1ZXM, solved to 1.87 Å resolution, reveals a nearly identical dexrazoxane binding pocket (Figure B.1C), validating the use of the ScTOP2 structure as a proxy for the human homologue. To determine whether to use the apo or holo structure, the binding pockets were compared. The OpenEye™ make_receptor program probes for the shape of a binding pocket and determines the volume of this shape. The location in which dexrazoxane binds is a relatively small, enclosed binding pocket. A comparison of the apo and holo structures revealed an increase in the size of the pocket as a result of dexrazoxane binding (Figure B.1D) and a displacement of surrounding water molecules. Based on these results, the holo structure, 1QZR was chosen as the target structure for the VLS. It was then prepared for docking in three different docking programs with the

Schrödinger® Protein Preparation Wizard, MGLTools, and OpenEye™ make_receptor.

Initial VLS with enriched ChemBridge™ Express Library identifies Schrödinger® Glide as best docking program for ScTOP2

The ChemBridge™ Express Library (www.chembridge.com) was chosen for the VLS based on its size, diversity, druglike qualities, and commercial availability. The library has over 460,000 compounds covering a diverse chemical space. The characteristics of dexrazoxane, as a known binder, were used to enrich this library via the OpenEye™ ROCS program, which filters compounds based on shape similarity. In certain cases, shape-screening alone can outperform traditional docking, making it an attractive step in a docking workflow [181]. The top 10^4 similar compounds were then prepared for docking through Schrödinger's LigPrep and OpenEye™'s OMEGA.

Using the prepared target protein structure, 1QZR, and the enriched ChemBridge™ Express Library, an initial VLS was conducted using Schrödinger® Glide, OpenEye™ fred, and AutoDock™ Vina, all set to run in medium precision mode.

While the use of a consensus score can outperform a single scoring function [182], the more common strategy is to find the program that produces the best results for the specific target in question, as no single docking program consistently outperforms the others [183]. Typically, this choice is made based on the program's ability to distinguish between hits and decoys in terms of docking rank: generally speaking, the hits should be well-ranked and the decoys should be poorly ranked.

In this case, only the ranking of ICRF-187 was used to determine which program best reproduced known binding results, as this is the only hit to have been crystallized in the binding site. Comparing the three programs' results, only Glide produced dexrazoxane as a hit within the top 100 hits, scoring 6th overall. For Vina, the top ranking ICRF-187 pose was ranked 267th, and for OpenEye, this rank was 658. Based on the promising results from the standard precision Glide screening, a more exhaustive screening was then conducted with Glide in extra-precision (XP) mode.

27 compounds selected from GlideXP tested for TOP2 inhibition via coupled ATPase assays

Using the more exhaustive GlideXP, the top scoring pose of dexrazoxane was ranked 3rd overall, further verifying that Glide's scoring function correctly selected for a known hit. For *in vitro* testing, 27 compounds were selected from the top 100 unique hits. Molecules were filtered out based on poor solubility, non-druglike moieties, and instability. Table 1 shows the compounds selected for screening.

These compounds were then tested *in vitro* against both ScTOP2 and hTOP2A using an NADH-coupled ATPase assay. In this assay, the consumption of ATP is directly correlated to the oxidation of NADH and the resultant loss of 340 nm fluorescence [184]. Thus ATPase activity is measured by the slope in the 340 nm absorbance reading. Based on these measurements, the IC_{50} for dexrazoxane against ScTOP2 was around 50 μM, while it was much more effective against hTOP2A with an IC_{50} of approximately 1 μM (Figure B.2).

Against the yeast target, the VLS compounds were ineffective in reducing ATPase activity (results not shown), but against the human homologue, they showed some inhibitory effect. While these results need to be repeated, several of the compounds showed some inhibition of hTOP2A ATPase activity at 5 and 50 μM concentrations (Figure B.3).

Future Directions

It is encouraging to see some measureable inhibitory activity from a number of the VLS compounds against the human target, hTOP2A. In virtual screening, it is not uncommon for initial hits to show limited inhibition. These initial compounds can then act as the basis for further lead optimization. Future directions would include developing a structure activity relationship (SAR), improved or constrained docking using information from these tested hits, and derivatization of initial leads to improve binding and/or drug-like characteristics.

Materials & Methods

Target selection. A comparison of the structure of ScTOP2 (PDB ID: 1QZR) with bound ICRF-187 to the apo structure of ScTOP2 (PDB ID: 1PVG) using the binding shape detection of OpenEye™ make_receptor led to the selection of 1QZR as the target structure, based on the cavity size: 748 Å^3 for 1QZR vs 518 Å^3 for 1PVG. The human and yeast structures, 1ZXM and 1QZR, respectively, were compared in PyMol using the color_by_conservation.py script [145].

Target preparation. The target was prepared with the Schrödinger® Protein Preparation Wizard, using the default values. This preparation removed bound waters, added hydrogens, optimized protonation states, flipped residues, and optimized the hydrogen bond network, among other functions. This prepared structure was then used to generate a grid for docking with Schrödinger® Glide, using the ligand to define the centroid of the docking volume. No constraints were chosen and no volumes were excluded. Six residues were chosen to be rotatable groups in the Glide docking: Thr27, Tyr28, and Tyr144 on both chains.

For use in AutoDock™ Vina, the Schrödinger® prepared protein was exported as a PDB file, imported into MGLTools (Scripps Research Institute®), processed to leave only polar hydrogens, stripped of the bound ligand, and then exported in pdbqt format. The docking volume was centered at (27.5, 33, 32.5), with a length of 13 Å in each dimension.

For use in OpenEye™, the prepared protein was imported into make_receptor, and the ligand was chosen to define the docking volume. The docking box was chosen to be 1490 $Å^3$, while the detected docking shape volume was 793 $Å^3$.

Ligand preparation. The ChemBridge™ Express library was filtered through Schrödinger® qikprop to eliminate non-druglike molecules. The next filtering step was accomplished with OpenEye™ ROCS: the top 10k compounds were selected based on dexrazoxane shape similarity. This filtered library was then put through LigPrep with default settings to desalt, generate tautomers, vary chiral centers, generate possible ionization states at pH values between 2 and 7 (via Epik), and minimize resulting structures using the OPLS_2005 force field. For OpenEye™, this prepared library was run through OMEGA to generate conformers, with a maximum of 2000 conformers per ligand. For Vina, each of these structures was converted to pdbqt format.

Docking. Using the prepared 1QZR structure and the prepared ligand library, all three programs were initially run in medium precision mode. Glide was run using Standard Precision (SP); AutoDock™ Vina was run with the exhaustiveness set to 4; and OpenEye™ was run with the Standard resolution for the exhaustive search, 1 Å translational step size, 1.5 Å rotational step size. AutoDock™ Vina is a script for docking one ligand into a target receptor. To adapt this to screening an entire ligand library, several bash scripts were used (Programs 1,2). For both Glide dockings, the settings were left at their default values.

Constructs and Protein preparation. Both full-length TOP2 proteins were encoded on a 12Ura-B plasmid with a GAL1 promoter from MacroLab (http://qb3.berkeley.edu/qb3/macrolab/) with an N-terminal His_6 affinity tag followed by a tobacco etch virus (TEV) protease cleavage site. Both hTOP2A and ScTOP2 were expressed in the *S. cerevisiae* BCY123 strain and purified via His-tag affinity chromatography and gel filtration chromatography, as in [170].

ATPase assays. The NADH-coupled reaction mixture contained a final concentration of 3.75 mM phosphoenolpyruvate (Sigma-Aldrich®), 4% v/v pyruvate kinase/lactic dehydrogenase enzyme mixture (Sigma-Aldrich®), 0.06 mg/mL BSA, 40 mM Tris pH 7.9, 3 mM TCEP, 10 mM $MgCl_2$, 0.132 mg/mL sheared salmon sperm DNA, 1% DMSO, 200 µM NADH, 100 nM TOP2 (yeast or human A isoform), and 5% glycerol. This mixture equilibrated at 30°C (ScTOP2) or 37°C (hTOP2A) for five minutes before the addition of ATP, for a final concentration of 4 mM ATP. For each condition, NADH fluorescence was monitored every 30 seconds for 2 h with a CLARIOstar® plate reader (BMG LABTECH). This raw fluorescence data was processed in MATLAB® to determine ATP consumption (M/s) (Programs 3-9). IC_{50} measurements were made with an adapted program from [185] (Program 8).

Tables

Table B.1: VLS compounds targeting eukaryotic TOP2

Comp.	IUPAC	CID	Structure	Mol Wt	cLogP
ICRF-187	4-[(2S)-2-(3,5-dioxopiperazin-1-yl)propyl]piperazine-2,6-dione	71384	[redacted]	268.3	-1.4
1	(3Z)-3-[(4-hydroxyanilino)methylidene]pyrrolidine-2,4-dione	1713548	[redacted]	218.2	-0.27
2	2-(3,5-dimethylpyridin-1-ium-1-yl)-1-phenylethanone;bromide	2840993	[redacted]	306.2	2.15
3	[1-[(E)-3-(2-nitrophenyl)prop-2-enyl]pyrrolidin-2-yl]methanol	5721669	[redacted]	262.3	2.22
4	3-[(4-hydroxyanilino)methylene]-2,4-pyrrolidinedione	45927571	[redacted]	210.4	3.17
5	N-(cyclohex-3-en-1-ylmethyl)-N,1-dimethylpiperidin-4-amine	2845571	[redacted]	222.4	2.26
6	1-[(E)-3-(2-nitrophenyl)prop-2-enyl]pyrrolidine	27267173	[redacted]	232.3	2.95
7	4-[(E)-3-(2-nitrophenyl)prop-2-enyl]thiomorpholine	783356		264.3	3.03
8	N-(2-piperidin-1-ylethyl)thiolan-3-amine	4107026	[redacted]	214.4	1.81
9	1-[2-(4-chlorophenoxy)ethyl]piperidine	756651	[redacted]	239.7	4.26
10	4-[2-(3-chlorophenoxy)ethyl]morpholine	644756	[redacted]	241.7	3.04
11	3,5-dimethyl-N-(pyridin-4-ylmethyl)aniline	791954	[redacted]	212.3	2.59
12	3,4-dimethoxy-N-pyridin-4-ylbenzamide	814761	[redacted]	258.3	1.74
13	N-(pyridin-4-ylmethyl)cyclohexanecarboxamide	823684	[redacted]	218.3	1.48
14	2-(2,4-dioxo-1,3-thiazolidin-5-yl)-N-(2-hydroxyphenyl)acetamide	2905088	[redacted]	266.3	0.51

15	N-(3-methylphenyl)pyridine-2-carbothioamide	797068	[redacted]	228.3	3.12
16	4-[2-(4-nitropyrazol-1-yl)ethyl]morpholine	2984635	[redacted]	226.2	0.521
17	1-[4-[2-(2,6-dimethylmorpholin-4-yl)ethoxy]-3-methoxyphenyl]ethanone;hydrochloride	2994535	[redacted]	343.9	-0.25
18	[2-(2-morpholin-4-ylethoxy)phenyl]methanol;hydrochloride	45595395	[redacted]	273.8	0.29
19	1-[2-(4-iodo-2-methylphenoxy)ethyl]-4-methylpiperazine;hydrochloride	6457337	[redacted]	396.7	3.21
20	1-[2-[2-(2,6-dimethylmorpholin-4-yl)ethoxy]phenyl]ethanone;hydrochloride	2996517	[redacted]	313.8	1.5
21	5-(2-morpholin-4-ylethyl)-1,3,4-thiadiazol-2-amine	3164058	[redacted]	214.3	-0.484
22	2-(3-methylpiperidin-1-yl)-N-(5-methyl-1,3,4-thiadiazol-2-yl)acetamide	3843682	[redacted]	254.4	1.777
23	2-(azepan-1-yl)-N-(1,3,4-thiadiazol-2-yl)acetamide	6469493	[redacted]	240.3	1.434
24	N-(2,5-dichlorophenyl)-2-(3-methylpiperidin-1-yl)acetamide	4563532	[redacted]	301.2	3.833
25	6-chloro-4-N-(pyridin-3-ylmethyl)pyrimidine-2,4-diamine	14257470	[redacted]	235.7	1.202
26	6-chloro-4-N-(2-morpholin-4-ylethyl)pyrimidine-2,4-diamine	45869411	[redacted]	257.7	1.467
27	5-[(4-methoxyanilino)methyl]-1,2-dihydro-1,2,4-triazol-3-one	45490304	[redacted]	220.2	0.397

Figures

Figure B.1. Eukaryotic TOP2 structure, inhibition, and target selection. A) Eukaryotic TOP2 is composed of two functional subunits: the ATPase domain and the DNA cleavage domain. Here, the G-segment is bound in the DNA gate of TOP2, with the T-segment in position to enter through the N-gate. B) The binding of dexrazoxane functions to "lock" the N-gate, stabilizing the dimer interface in a closed position. C) A comparison of ScTOP2 and hTOP2A reveals a high degree of conservation in the dexrazoxane binding pocket, as visualized here using the PyMol color_by_conservation script. Red indicates well-conserved while blue indicates some divergence. D) OpenEye™ make_receptor detects the shape of the binding pocket and gives a volume estimation. Here, the volume of the holo ScTOP2 binding pocket is ~200 Å larger than the apo structure.

Figure B.1

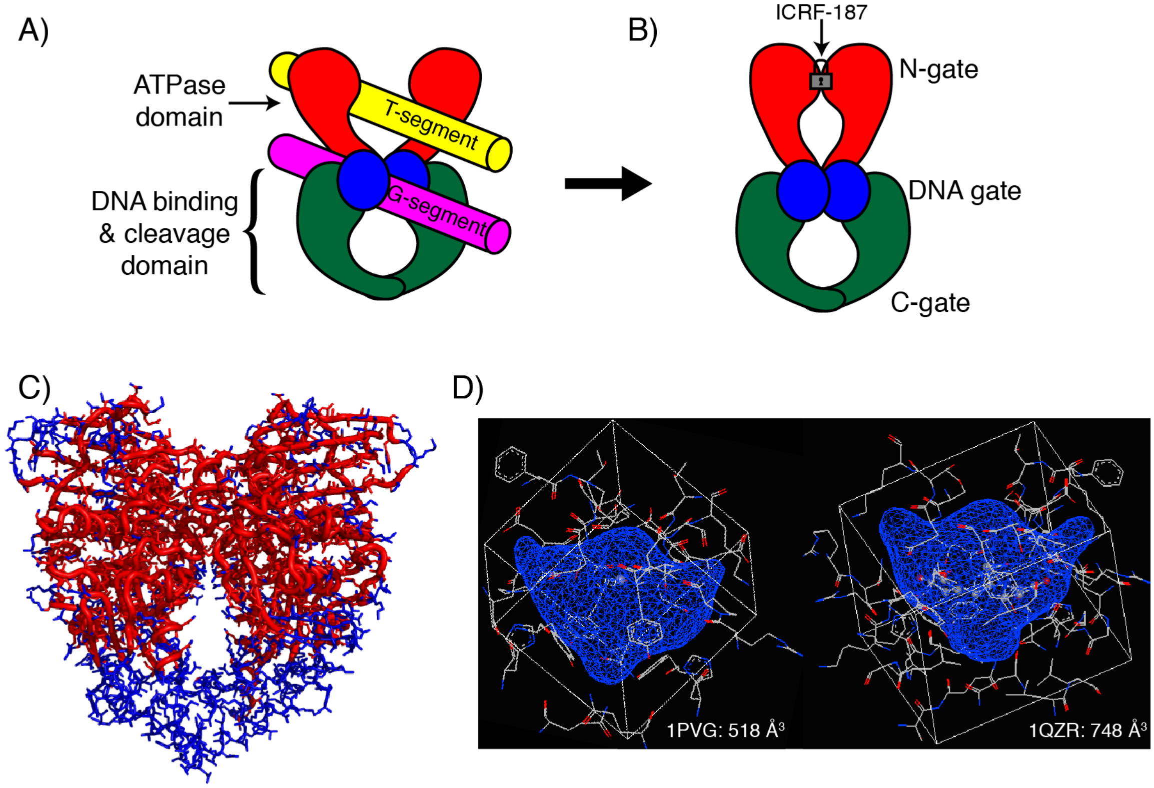

A)
ATPase domain
T-segment
DNA binding & cleavage domain
G-segment
B)
ICRF-187
N-gate
DNA gate
C-gate
C)
D)
1PVG: 518 Å3
1QZR: 748 Å3

Figure B.2. Dexrazoxane inhibition of eukaryotic TOP2 ATPase activity. A) Raw 340 nm fluorescence readings for 2-fold titration series of dexrazoxane inhibiting hTOP2A. B) Slopes from linear regression of raw data analyzed to establish *in vitro* IC_{50} against hTOP2A. C) As in (A), but for ScTOP2. D) As in (B), but for ScTOP2.

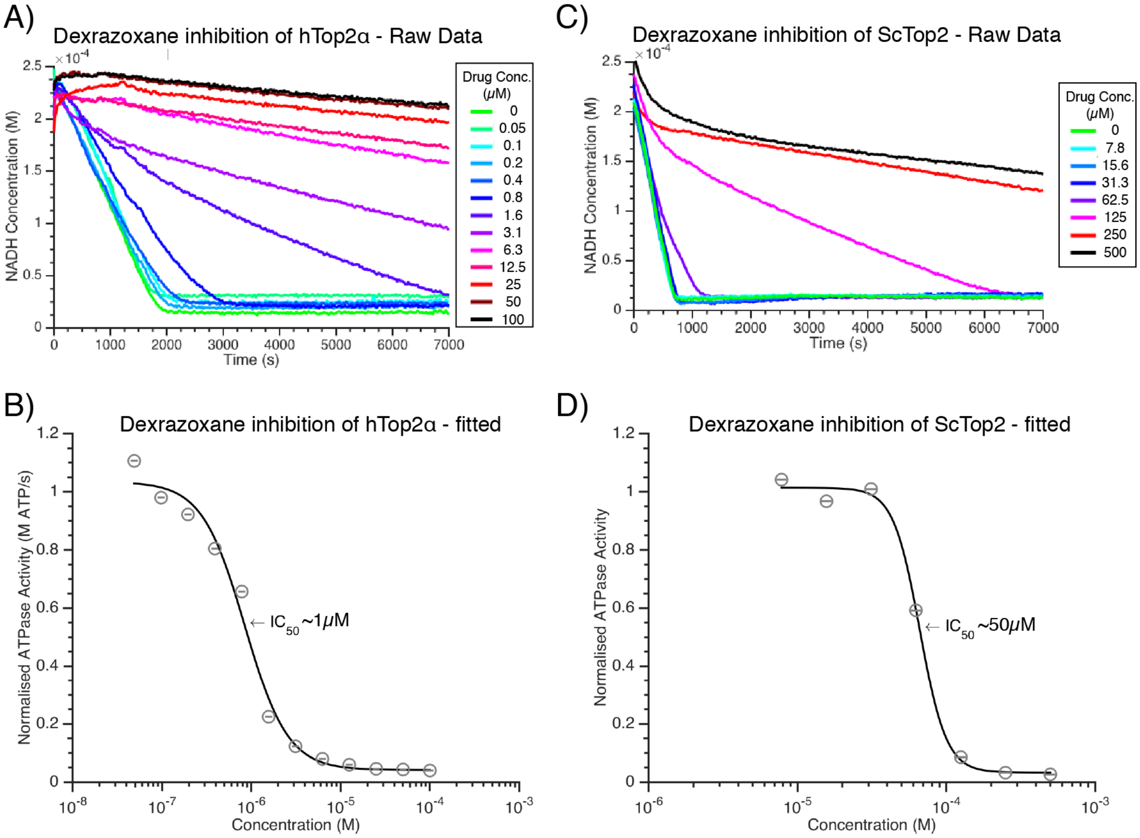
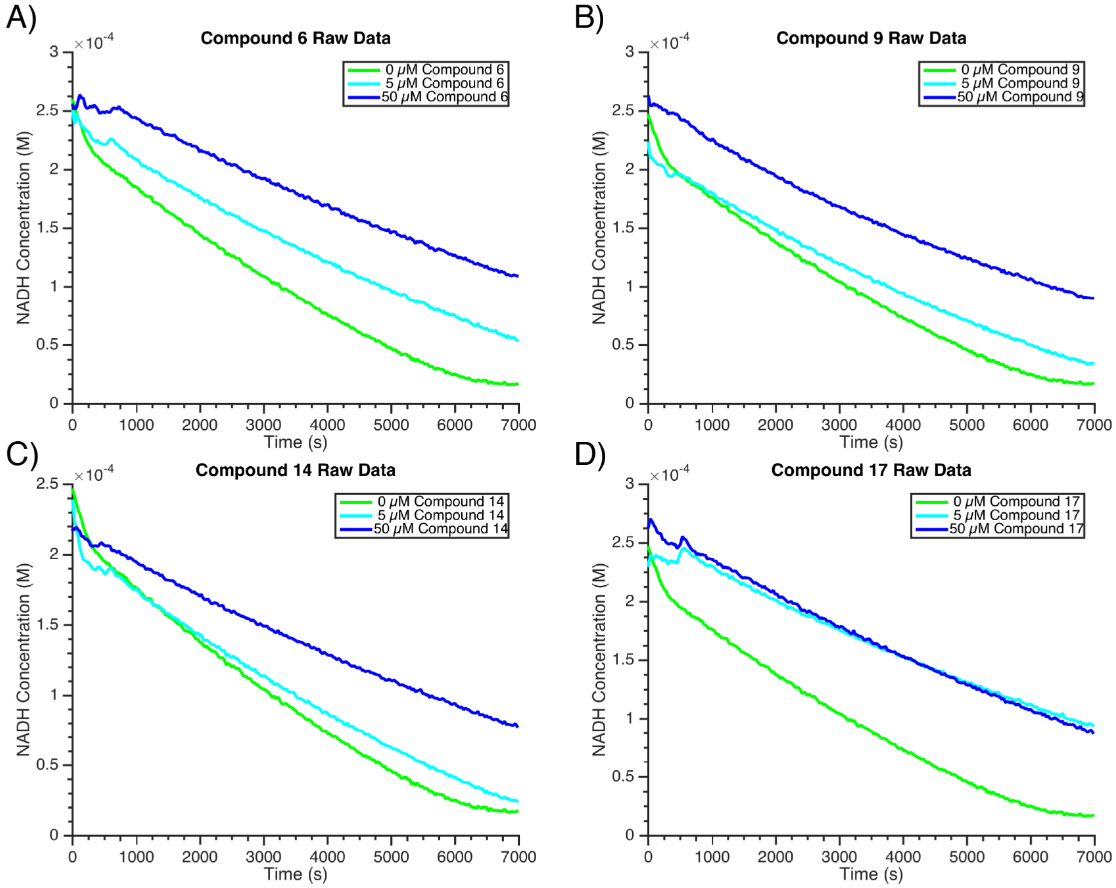


Figure B.3. VLS compound inhibition of hTOP2A ATPase activity. 340 nm fluorescence readings for NADH-coupled ATPase assays in the presence of 0, 5, and 50 μM of (A) Compound 6, (B) Compound 9, (C) Compound 14, and (D) Compound 17 shows some inhibition of ATPase activity.



Programs

Program 1: vina_library.sh

```
#!/bin/bash -x

#  vina_library.sh
#  A script to run vina on a large library. Takes in the ligand
root directory
#
#  Created by Alexia Miller on 2/13/15.
#

# User must specify a ligand directory
if [ "$1" != "" ]; then
    ligDir="$1"
else
    echo "Please specify a ligand directory!"
    exit 1
fi

# This is the ligand directory name, without the full path, e.g.
CBExp01qt
prefix=`basename $ligDir`

# Making new folders for the results and the log files
resPath="${prefix}_Results"
logPath="${prefix}_Logs"
mkdir $resPath
mkdir $logPath

# Take out trailing slash in ligDir
ligDir=$(echo "$ligDir"|sed 's/\/$//g')

source /programs/sbgrid.shrc

# For all the ligands (pdbqt files) in this directory
for f in ${ligDir}/*.pdbqt; do
    # b is the ligand name without the file extension
    b=`basename $f .pdbqt`
    # Reports what ligand we're working on
    echo "Processing ligand $b"
    vina --config conf.txt --ligand $f --out
${resPath}/${b}_out.pdbqt --log ${logPath}/${b}_log.txt
done
```

Program 2: vina_to_pdb.sh

```bash
#!/bin/bash

#  vina_to_pdb.sh
#  A script to convert vina hits to pdb files and copy them to a
#  tophits folder with rank indicated in the new filename
#
#  Created by Alexia Miller on 3/24/15.
#

# User must specify the filename for the Vina hits
echo "Please specify the filename containing the output of
vina_screen_get_top.py. Then press [ENTER]."
read inFile

# If nothing is entered, then exit the program
if [ -z "$inFile" ]; then
    echo -n "No filename specified."
    exit 1
fi

# Make directory for all of the tophits in the file
mkdir tophits

source /programs/sbgrid.shrc

# Remove first line from file
sed '/Found/d' ./$inFile >temp_top.txt

rank=1

# Translate spaces to returns in the file
for i in `tr ' ' '\n' <temp_top.txt`; do
    # isolate original directory
    allHitsDir="${i%/*}"
    # filename without directory or extension
    hitname=`basename $i .pdbqt`
    # make new name with hit number
    rankname="${hitname}_rank${rank}"
    # convert to pdb and move to tophits directory
    obabel -ipdbqt $i -opdb -O "tophits/${rankname}.pdb"
    # index the rank
    rank=$(( rank + 1 ))
done

rm temp_top.txt
```

Program 3: myLabs.m

```
function myLabs=concRanges(myConcs,iD)
%% Concentration manipulations
% Finding a nice concentration range for plotting
if myConcs(1)<10^-3
    myConcLab=myConcs*10^6;
    concRange='μM';
elseif myConcs(1)<1
    myConcLab=myConcs*10^3;
    concRange='mM';
else
    myConcLab=myConcs;
    concRange='M';
end
% Concentration labels
myLabs=strcat(num2str(myConcLab'),[' ' concRange ' ' iD]);
```

Program 4: varAssign.m

```
% A function that parses the atpase assay output from Clariostar
% For use with nadhCSV, plotATPCSV, and getSlopes
function
[myData,timePts]=varAssign(dirName,fileName,rowOff,timeStep,myInd)
%% Variable assignment
% Finding the number of data columns (the number of reads)
delimiter = ',';
fid = fopen([dirName fileName],'rt');
i=1;
while i<=rowOff+1
    tLines = fgetl(fid);
    i=i+1;
end
numCols = numel(strfind(tLines,delimiter));
fclose(fid);

% Reading in sampleID number and raw absorbance values from the csv
file
fid=fopen([dirName fileName]);
allData=textscan(fid,['%*s %*s Sample X%d',repmat('%f',[1,numCols-
3])],...
    'CollectOutput',1','HeaderLines',rowOff,'Delimiter',',');
fclose(fid);
% All of the sample ID numbers
unsortedID=allData{:,1};
% All of the data for each sample
unsortedRaw=allData{:,2};

% Sorting the data based on the ID number
[~,IX] = sort(unsortedID);
rawData = unsortedRaw(IX,:);

% Assigning the data appropriately
myData=rawData(myInd,:);

% These are the time points, x axis data, in seconds
timePts=0:timeStep:(timeStep*size(rawData,2)-1);
```

Program 5: nadhCSV.m

```
% This is a function which reads in a csv file with the results of a
% coupled atpase assay and returns the NADH standard curves
% dirName=path to the directory containing your file;
% fileName=csv results file name;
% rowOff=row offset (after all the header lines);
% timeStep=time in seconds between each plate read
% nInd=indices of the NADH standards
% nConcs=concentrations of NADH in each of those standards, given in M

function
[nSlope,nIntcpt]=nadhCSV(dirName,fileName,rowOff,timeStep,nInd,nConcs,s
aveID)
%% Variable assignment
[nData,timePts]=varAssign(dirName,fileName,rowOff,timeStep,nInd);

%% Concentration manipulations
% If no concentration is given, this is the default (top is 200 μM)
if isempty(nConcs)
    nConcs=[200,100,50,0]*10^-6;
end

% Determining if concentrations are low-high and then reversing to
high-low
if nData(1,1)<nData(end,1)
    nConcs=wrev(nConcs);
end

% Labels for plots including the appropriate concentrations
nLabs=concRanges(nConcs,'NADH');

%% NADH Raw Data
% New figure, hfig is the "handle" for this figure
hfig=figure;
% Plotting raw data for the NADH titration
plot(timePts,nData','LineWidth',3)
% Plot labels
legend(nLabs)
title('NADH Raw Data')
xlabel('Time (s)')
ylabel('340 nm Absorbance')
% Plot style
set(gca,'FontSize',16)
set(gca, 'FontName', 'Helvetica')
set(gca, 'Box', 'off', 'TickDir', 'out', 'TickLength', [.02 .02], ...
    'XMinorTick', 'on', 'YMinorTick', 'on','LineWidth',2)

% Filename includes directory and name of figure
filename=[dirName saveID '_raw_data.png'];
% Saving the image
print(hfig,'-dpng','-r300',filename)

%% NADH Standard calculations
% Calculating average absorbances - just averages the whole curve
nMeanAbs=mean(nData,2);
```

```matlab
% Fitting a linear curve to the four absorbances
nLM=fitlm(nConcs,nMeanAbs);
% Extracting the slope from the fit
nSlope=nLM.Coefficients.Estimate(2);
% Extracting the intercept from the fit
nIntcpt=nLM.Coefficients.Estimate(1);

% New figure
hfig=figure;
% Plotting the average absorbances
scatter(nConcs,nMeanAbs,200,'filled')
hold on
% Plotting the linear fit on top of the data
concX=linspace(0,max(nConcs),300);
absY=nSlope*concX+nIntcpt;
plot(concX,absY,'LineWidth',3,'LineStyle','--','Color','k')
hold off

% Plot labels
title('NADH Standard Curve')
xlabel('NADH Concentration (M)')
ylabel('Mean 340 nm Absorbance')
% Plot style
set(gca,'FontSize',16)
set(gca, 'FontName', 'Helvetica')
set(gca, 'Box', 'off', 'TickDir', 'out', 'TickLength', [.02 .02], ...
    'XMinorTick', 'on', 'YMinorTick', 'on','LineWidth',2)

% Filename includes directory and name of figure
filename=[dirName saveID '_standard_curve.png'];
% Saving the image
print(hfig,'-dpng','-r300',filename)

end
```

Program 6: plotATPCSV.m

```
% A program that plots the data from an atpase assay
% must first use nadhCSV
function
plotATPCSV(dirName,fileName,rowOff,dataID,timeStep,myInd,myConcs,nSlope
,nIntcpt,saveID)
%% Variable assignment
[myData,timePts]=varAssign(dirName,fileName,rowOff,timeStep,myInd);

%% Concentration manipulations
myLabs=concRanges(myConcs,saveID);

%% Converting and plotting raw data

% Converting from absorbance to NADH concentration
dataNorm=(myData-nIntcpt)./nSlope;

% Defining colors for the graph
ColorSet = varycolor(length(myInd));
% hfig is the figure handle for a new figure
hfig=figure;
hold on
% For all the samples, plot on the same figure by using a for loop
for i=1:length(myInd)
    % Plot all the raw data with different colored lines
    plot(timePts, dataNorm(i,:)','Color',ColorSet(i,:),'LineWidth',3)
    %scatter(timePts,
dataNorm(i,:)',75,'Marker','s','MarkerFaceColor',...
    %    ColorSet(i,:),'MarkerEdgeColor',ColorSet(i,:))
end
hold off
% Plot labels
legend(myLabs)
title([saveID ' Raw Data'])
xlabel('Time (s)')
ylabel('NADH Concentration (M)')
% Plot style settings
set(gca,'FontSize',16)
set(gca, 'FontName', 'Helvetica')
set(gca, 'Box', 'off', 'TickDir', 'out', 'TickLength', [.02 .02], ...
    'XMinorTick', 'on', 'YMinorTick', 'on','LineWidth',2)

% Filename includes directory and name of figure
filename=[dirName dataID '_raw.png'];
% Saving the image
print(hfig,'-dpng','-r300',filename)

end
```

Program 7: getSlopes.m

```
% plots as in plotATPCSV.m, but then allows user to select linear
regions of each plot
% outputs plot of slopes
function
dataSlopes=getSlopes(dirName,fileName,rowOff,dataName,timeStep,myInd,my
Concs,nSlope,nIntcpt,iD)
%% Variable assignment
[myData,timePts]=varAssign(dirName,fileName,rowOff,timeStep,myInd);

%% Concentration manipulations
myLabs=concRanges(myConcs,iD);

%% Converting and plotting raw data
% Converting from absorpance to NADH concentration
dataNorm=(myData-nIntcpt)./nSlope;

% Defining colors for the graph
ColorSet = varycolor(length(myInd));

% Defining the linear regions of each curve
for i=1:length(myInd)
    figure
    plot(timePts, dataNorm(i,:)','Color',ColorSet(i,:))
    selPts=selectdata('selectionmode','rect');
    begPts(i)=selPts(1);
    endPts(i)=selPts(end);
end
close all

%% Plotting data
hfig=figure;
hold on
for i=1:length(myInd)
    plot(timePts, dataNorm(i,:)','Color',ColorSet(i,:),'LineWidth',3)
end

% Plot labels
legend(myLabs)
title([iD ' Raw Data'])
xlabel('Time (s)')
ylabel('NADH Concentration (M)')
% Plot style
set(gca,'FontSize',16)
set(gca, 'FontName', 'Helvetica')
set(gca, 'Box', 'off', 'TickDir', 'out', 'TickLength', [.02 .02], ...
    'XMinorTick', 'on', 'YMinorTick', 'on','LineWidth',2)

% Filename includes directory and name of figure
filename=[dirName dataName '_unfitted.png'];
% Saving the image
print(hfig,'-dpng','-r300',filename)

%% Fitting to the linear regions and plotting the fit
for i=1:length(myInd)
```

```
dataLM=fitlm(timePts(begPts(i):endPts(i)),dataNorm(i,begPts(i):endPts(i
)));
    dataIntcpts(i)=dataLM.Coefficients.Estimate(1);
    dataSlopes(i)=dataLM.Coefficients.Estimate(2);
    dataLine=dataIntcpts(i)+dataSlopes(i)*timePts(begPts(i):endPts(i));

plot(timePts(begPts(i):endPts(i)),dataLine,'Color','k','LineStyle',...
        '--','LineWidth',3)
end
legend(myLabs)
title([iD ' Linear Fitted Data'])
xlabel('Time (s)')
ylabel('NADH Concentration (M)')

set(gca,'FontSize',16)
set(gca, 'FontName', 'Helvetica')
set(gca, 'Box', 'off', 'TickDir', 'out', 'TickLength', [.02 .02], ...
    'XMinorTick', 'on', 'YMinorTick', 'on','LineWidth',2)
hold off

% Filename includes directory and name of figure
filename=[dirName dataName '_fitted.png'];
% Saving the image
print(hfig,'-dpng','-r300',filename)

%%
hfig=figure;
scatter(myConcs,-1*dataSlopes,150,'*')
hold on
line(myConcs,-1*dataSlopes,'LineWidth',3)
title([iD ' Titration'])
xlabel([iD ' Concentration (M)'])
ylabel('Velocity (M ATP/s)')

set(gca,'xscale','log')
set(gca,'FontSize',16)
set(gca, 'FontName', 'Helvetica')
set(gca, 'Box', 'off', 'TickDir', 'out', 'TickLength', [.02 .02], ...
    'XMinorTick', 'on', 'YMinorTick', 'on','LineWidth',2)

% Reference line for a 0 drug control
if sum(myConcs==0)>0
    ind=find(myConcs==0,1);
    xl=xlim;
    plot(xl,[-1*dataSlopes(ind) -1*dataSlopes(ind)],'LineWidth',3,...
        'LineStyle','--','Color','k')
end

hold off

% Filename includes directory and name of figure
filename=[dirName dataName '_slopes.png'];
% Saving the image
print(hfig,'-dpng','-r300',filename)
end
```

Program 8: doseResponseATPase.m

NOTE: This function is essentially the doseResponse function [185] with a few edits to make it appropriate for nadh-coupled atpase activity assays.

Copyright notice:

```
Copyright (c) 2011, Ritchie Smith
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are
permitted provided that the following conditions are met:

    * Redistributions of source code must retain the above copyright notice, this list of
conditions and the following disclaimer.
    * Redistributions in binary form must reproduce the above copyright notice, this list
of conditions and the following disclaimer in the documentation and/or other materials
provided with the distribution

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY
EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF
MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL
THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL,
SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT
OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION)
HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR
TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS
SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.
```

```
function
[hillCoeff,ic50]=doseResponseATPase(dirName,myConcs,mySlopes,dataID,sav
eID)

% DOSERESPONSE     Computes the Hill Coefficient and EC50 of a
% myConcs/mySlopes relationship given two vectors describing the
myConcss and
% mySlopess. A semilog graph is plotted illustrating the relationship,
upon
% which the man and standard error of the mySlopes to each myConcs
level is
% plotted along with the fitted Hill Equation sigmoid. The EC50 is also
% labelled. Requires nlinfit from statistics toolbox.
%
%   Example
%       d=[3.75 3.75 3.75 3.75 3.75 7.5 7.5 7.5 7.5 15 15 15 15 15 60
60 60 60]
%       r=[107 91 99 124 100 96 92 133 119 84 66 86 106 91 52 37 10 69]
%       [hill ic50]=myConcsResponse(d,r)
%
%   Inputs
%       Two vectors of the same length, the first containing the
myConcs and
%       the second the mySlopes. If myConcss of 0 are contained in the
data,
%       these are used as control values and the data are normalised by
%       their mean value.
%
%   Notes
%       This function was written to rapidly produce simple
myConcs/mySlopes curves
%       and EC50s for publication.
```

```
%deal with 0 dosage by using it to normalise the results.
normalised=0;
if (sum(myConcs(:)==0)>0)
    %compute mean control mySlopes
    controlResponse=mean(mySlopes(myConcs==0));
    %remove controls from myConcs/mySlopes curve
    mySlopes=mySlopes(myConcs~=0)/controlResponse;
    myConcs=myConcs(myConcs~=0);
    normalised=1;
end

%hill equation sigmoid
sigmoid=@(beta,x)beta(1)+(beta(2)-beta(1))./(1+(x/beta(3)).^beta(4));

%calculate some rough guesses for initial parameters
minResponse=min(mySlopes);
maxResponse=max(mySlopes);
midResponse=mean([minResponse maxResponse]);
minDose=min(myConcs);
maxDose=max(myConcs);

%fit the curve and compute the values
[coeffs,r,J]=nlinfit(myConcs,mySlopes,sigmoid,[minResponse maxResponse
midResponse 1]);

ic50=coeffs(3);
hillCoeff=coeffs(4);

%plot the fitted sigmoid
hfig=figure;
xpoints=logspace(log10(minDose),log10(maxDose),1000);
semilogx(xpoints,sigmoid(coeffs,xpoints),'Color',[0 0 0],'LineWidth',2)
hold on

%notate the IC50
text(ic50,mean([coeffs(1) coeffs(2)]),[' \leftarrow '
sprintf('IC_{50}=%0.2g',ic50)],'FontSize',16);

%plot mean mySlopes for each myConcs with standard error
myConcss=unique(myConcs);
meanResponse=zeros(1,length(myConcss));
stdErrResponse=zeros(1,length(myConcss));
for i=1:length(myConcss)
    mySlopess=mySlopes(myConcs==myConcss(i));
    meanResponse(i)=mean(mySlopess);
    stdErrResponse(i)=std(mySlopess)/sqrt(length(mySlopess));
    %stdErrResponse(i)=std(mySlopess);
end

errorbar(myConcss,meanResponse,stdErrResponse,'o','Color',[0.5 0.5
0.5],'LineWidth',2,'MarkerSize',12)

% Label plot
title([dataID ' Titration'])
xlabel('Concentration (M)','FontSize',16);
```

```
if normalised
    ylabel('Normalised ATPase Activity','FontSize',16);
else
    ylabel('ATPase Activity (M ATP/s)','FontSize',16);
end
set(gca,'FontSize',16);
set(gca, 'FontName', 'Helvetica')
set(gca, 'Box', 'off', 'TickDir', 'out', 'TickLength', [.02 .02], ...
    'XMinorTick', 'on', 'YMinorTick', 'on','LineWidth',2)

% Filename includes directory and name of figure
filename=[dirName saveID '_ic50.png'];
% Saving the image
print(hfig,'-dpng','-r300',filename)

hold off;
```

Program 9: coupledATPaseAnalysisTemplate.m

```
%% Experimental data location and details
% The directory where you have put your csv file
dirName='~/Documents/BoschLab/Berger/Data/ATPase/15-09-
18_dexrazoxane/';
% The name of your csv file
fileName='15-09-18_ScTopoII_ATPase_dex_control.txt';
% This is the time in seconds between each read
% In this example, reads were taken every 20 seconds
timeStep=20;

%% NADH standard calculations
% The sample IDs corresponding to your NADH titration
% In this example, the first four samples were the NADH titration
nInd=1:4;

% The NADH concentrations of your titration. Concentrations are in M!
% In this example, the concentrations are 200, 100, 50, and 0
micromolar
nConcs=[200 100 50 0]*10^-6;

% The number of header rows (usually 7)
rowOff=7;
% Prefix for saving the output images
saveID='NADH';

% The outputs of this function are the results of fitting the NADH
standard
% curve: the slope and intercept of the line
[nSlope,nIntcpt]=nadhCSV(dirName,fileName,rowOff,timeStep,nInd,nConcs,s
aveID);

%% Plotting data for a titration series
% The sample IDs corresponding to the samples you would like to plot
% In this example, the 5th through 12th samples are the ones to plot
myInd=5:12;
% These are the concentrations of the samples you want to plot (in M)
% In this example, I have a dilution series ranging from 100 to 0
micromolar
myConcs=[100 50 25 12.5 6.25 3.125 1.5625 0]*10^-6;

% % Another way to calculate a dilution series
% % The high concentration is 100 micromolar
% highConc=100*10^-6;
% % If you are doing seven two-fold dilutions with a 0 control at the
end
% titrationSteps=0:1:6;
% twoFolds=2.^titrationSteps;
% dilutSeries=[twoFolds inf];
% myConcs=highConc./dilutSeries;

% This is the name of the drug/ATP/DNA/enzyme that you are titrating.
It
% will be used in the title and figure legend
% In this example, I am titrating dexrazoxane
```

```
dataID='Dexrazoxane';
% Prefix for saving output images
saveID2='Dex';

% This will produce a nice plot labeled appropriately
plotATPCSV(dirName,fileName,rowOff,saveID2,timeStep,myInd,myConcs,nSlop
e,nIntcpt,dataID)
%% Getting slopes
% You can use this function to select the linear portion of each of
your
% curves
mySlopes=getSlopes(dirName,fileName,rowOff,saveID2,timeStep,myInd,myCon
cs,nSlope,nIntcpt,dataID)

%% Plotting IC50
% If you have performed a drug titration, this function will output the
% IC50 and the Hill coefficient for that titration.
% You MUST run getSlopes before trying to plot the IC50.
[hillCoeff,ic50]=doseResponseATPase(dirName,myConcs,mySlopes,dataID,sav
eID2)
```

BIBLIOGRAPHY

1. Kappe, S.H.I., Vaughan, A.M., Boddey, J.A., and Cowman, A.F. (2010) That was then but this is now: malaria research in the time of an eradication agenda. *Science* **328** (5980), 862–6.
2. Schwartz, L., Brown, G. V, Genton, B., and Moorthy, V.S. (2012) A review of malaria vaccine clinical projects based on the WHO rainbow table. *Malar. J.* **11** (1), 11.
3. Bhatt, S., Weiss, D.J., Cameron, E., et al. (2015) The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature* **526** (7572), 207–211.
4. Greenwood, B.M. (2015) Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: Final results of a phase 3, individually randomised, controlled trial. *Lancet* **386** (9988), 31–45.
5. Sibley, C.H. (2015) Understanding artemisinin resistance. *Science* **347** (6220), 373–374.
6. Mizushima, N., Yoshimori, T., and Ohsumi, Y. (2011) The Role of Atg Proteins in Autophagosome Formation. *Annu. Rev. Cell Dev. Biol.* **27** (1), 107–132.
7. Reggiori, F., and Klionsky, D.J. (2002) Autophagy in the Eukaryotic Cell. *Eukaryot. Cell* **1** (1), 11–21.
8. Mariño, G., Madeo, F., and Kroemer, G. (2011) Autophagy for tissue homeostasis and neuroprotection. *Curr. Opin. Cell Biol.* **23** (2), 198–

206.

9. Ryter, S.W., Cloonan, S.M., and Choi, A.M.K. (2013) Autophagy: A critical regulator of cellular metabolism and homeostasis. *Mol. Cells* **36** (1), 7–16.

10. Russell, R.C., Yuan, H.-X., and Guan, K.-L. (2014) Autophagy regulation by nutrient signaling. *Cell Res.* **24** (1), 42–57.

11. Kroemer, G., Mariño, G., and Levine, B. (2010) Autophagy and the Integrated Stress Response. *Mol. Cell* **40** (2), 280–293.

12. Murrow, L., and Debnath, J. (2012) Autophagy as a Stress-Response and Quality-Control Mechanism: Implications for Cell Injury and Human Disease. *Annu. Rev. Pathol. Mech. Dis.* **8** (1), 105–137.

13. Mizushima, N., and Levine, B. (2010) Autophagy in mammalian development and differentiation. *Nat. Cell Biol.* **12** (9), 823–830.

14. Rubinsztein, D.C., Codogno, P., and Levine, B. (2012) Autophagy modulation as a potential therapeutic target for diverse diseases. *Nat. Rev. Drug Discov.* **11** (9), 709–730.

15. Fuchs, Y., and Steller, H. (2015) Live to die another way: modes of programmed cell death and the signals emanating from dying cells. *Nat. Rev. Mol. Cell Biol.* **16** (6), 329–344.

16. Klionsky, D.J., Baehrecke, E.H., Brumell, J.H., et al. (2011) A comprehensive glossary of autophagy-related molecules and processes (2nd ed.). *Autophagy* **7** (11), 1273–1294.

17. Hain, A.U.P., and Bosch, J. (2013) Autophagy in Plasmodium, a multifunctional pathway? *Comput. Struct. Biotechnol. J.* **8** (11), e201308002.

18. Rubinsztein, D.C., Shpilka, T., and Elazar, Z. (2012) Mechanisms of autophagosome biogenesis. *Curr. Biol.* **22** (1), R29–R34.

19. Kamada, Y., Yoshino, K., Kondo, C., et al. (2010) Tor directly controls the Atg1 kinase complex to regulate autophagy. *Mol. Cell. Biol.* **30** (4), 1049–1058.

20. Stjepanovic, G., Davies, C.W., Stanley, R.E., et al. (2014) Assembly and dynamics of the autophagy-initiating Atg1 complex. *Proc. Natl. Acad. Sci.* **111** (35), 12793–12798.

21. Jaber, N., Dou, Z., Chen, J.-S., et al. (2012) Class III PI3K Vps34 plays an essential role in autophagy and in heart and liver function. *Proc. Natl. Acad. Sci.* **109** (6), 2003–2008.

22. Araki, Y., Ku, W.C., Akioka, M., et al. (2013) Atg38 is required for autophagy-specific phosphatidylinositol 3-kinase complex integrity. *J. Cell Biol.* **203** (2), 299–313.

23. Blommaart, E.F., Krause, U., Schellens, J.P., et al. (1997) The phosphatidylinositol 3-kinase inhibitors wortmannin and LY294002 inhibit autophagy in isolated rat hepatocytes. *Eur. J. Biochem.* **243**, 240–246.

24. Suzuki, K., Akioka, M., Kondo-Kakuta, C., et al. (2013) Fine mapping of autophagy-related proteins during autophagosome

formation in Saccharomyces cerevisiae. *J. Cell Sci.* **126** (Pt 11), 2534–44.

25. Yamamoto, H., Kakuta, S., Watanabe, T.M., et al. (2012) Atg9 vesicles are an important membrane source during early steps of autophagosome formation. *J. Cell Biol.* **198** (2), 219–233.
26. Xie, Z., Nair, U., and Klionsky, D.J. (2008) Atg8 Controls Phagophore Expansion during Autophagosome Formation. *Mol. Biol. Cell* **19**, 3290–3298.
27. Nakatogawa, H., Ishii, J., Asai, E., and Ohsumi, Y. (2012) Atg4 recycles inappropriately lipidated Atg8 to promote autophagosome biogenesis. *Autophagy* **8** (2), 177–186.
28. Geng, J., and Klionsky, D.J. (2008) The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. “Protein modifications: beyond the usual suspects” review series. *EMBO Rep.* **9** (9), 859–864.
29. Sakoh-Nakatogawa, M., Matoba, K., Asai, E., et al. (2013) Atg12–Atg5 conjugate enhances E2 activity of Atg3 by rearranging its catalytic site. *Nat. Struct. Mol. Biol.* **20** (4), 433–439.
30. Hanada, T., Noda, N.N., Satomi, Y., et al. (2007) The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy. *J. Biol. Chem.* **282** (52), 37298–37302.
31. Pappas, G., Roussos, N., and Falagas, M.E. (2009) Toxoplasmosis snapshots: Global status of Toxoplasma gondii seroprevalence and

implications for pregnancy and congenital toxoplasmosis. *Int. J. Parasitol.* **39** (12), 1385–1394.

32. Baldursson, S., and Karanis, P. (2011) Waterborne transmission of protozoan parasites: review of worldwide outbreaks - an update 2004-2010. *Water Res.* **45** (20), 6603–6614.
33. Dubey, J.P., and Schares, G. (2011) Neosporosis in animals - The last five years. *Vet. Parasitol.* **180** (1–2), 90–108.
34. Chapman, H.D. (2014) Milestones in avian coccidiosis research: a review. *Poult. Sci.* **93** (3), 501–11.
35. Reichel, M.P., Alejandra Ayanegui-Alcérreca, M., Gondim, L.F.P., and Ellis, J.T. (2013) What is the global economic impact of Neospora caninum in cattle - The billion dollar question. *Int. J. Parasitol.* **43** (2), 133–142.
36. Tun, K.M., Imwong, M., Lwin, K.M., et al. (2015) Spread of artemisinin-resistant Plasmodium falciparum in Myanmar: a cross-sectional survey of the K13 molecular marker. *Lancet Infect. Dis.* **15** (4), 415–421.
37. Bargieri, D., Lagal, V., Andenmatten, N., et al. (2014) Host Cell Invasion by Apicomplexan Parasites: The Junction Conundrum. *PLoS Pathog.* **10** (9), e1004273.
38. van Dooren, G.G., and Striepen, B. (2013) The algal past and parasite present of the apicoplast. *Annu. Rev. Microbiol.* **67**, 271–89.
39. Aurrecoechea, C., Barreto, A., Brestelli, J., et al. (2013) EuPathDB:

the eukaryotic pathogen database. *Nucleic Acids Res* **41** (D1), D684-91.

40. Wasmuth, J., Daub, J., Peregrín-Alvarez, J.M., et al. (2009) The origins of apicomplexan sequence innovation. *Genome Res.* **19** (7), 1202–1213.
41. Rigden, D.J., Michels, P.A., and Ginger, M.L. (2009) Autophagy in protists. *Autophagy* **5** (6), 784–794.
42. Duszenko, M., Ginger, M.L., Brennand, A., et al. (2011) Autophagy in protists. *Autophagy* **7** (2), 127–158.
43. Brennand, A., Gualdrón-López, M., Coppens, I., et al. (2011) Autophagy in parasitic protists: Unique features and drug targets. *Mol. Biochem. Parasitol.* **177** (2), 83–99.
44. Navale, R., Atul, Allanki, A.D., and Sijwali, P.S. (2014) Characterization of the Autophagy Marker Protein Atg8 Reveals Atypical Features of Autophagy in Plasmodium falciparum. *PLoS One* **9** (11), 1–36.
45. Tomlins, A.M., Ben-Rached, F., Williams, R.A., et al. (2013) Plasmodium falciparum ATG8 implicated in both autophagy and apicoplast formation. *Autophagy* **9** (10), 1540–52.
46. Cervantes, S., Bunnik, E.M., Saraf, A., et al. (2014) The multifunctional autophagy pathway in the human malaria parasite, Plasmodium falciparum. *Autophagy* **10** (1), 80–92.
47. Walker, D.M., Mahfooz, N., Kemme, K. a., et al. (2013) Plasmodium

falciparum Erythrocytic Stage Parasites Require the Putative Autophagy Protein PfAtg7 for Normal Growth. *PLoS One* **8** (6), e67047.

48. Besteiro, S., Brooks, C.F., Striepen, B., and Dubremetz, J.F. (2011) Autophagy protein Atg3 is essential for maintaining mitochondrial integrity and for normal intracellular development of toxoplasma gondii tachyzoites. *PLoS Pathog.* **7** (12), e1002416.
49. Besteiro, S. (2012) Which roles for autophagy in Toxoplasma gondii and related apicomplexan parasites? *Mol. Biochem. Parasitol.* **184** (1), 1–8.
50. Kong-Hap, M.A., Mouammine, A., Daher, W., et al. (2013) Regulation of ATG8 membrane association by ATG4 in the parasitic protist Toxoplasma gondii. *Autophagy* **9** (9), 1334–48.
51. Kirisako, T., Ichimura, Y., Okada, H., et al. (2000) The reversible modification regulates the membrane-binding state of Apg8/Aut7 essential for autophagy and the cytoplasm to vacuole targeting pathway. *J. Cell Biol.* **151** (2), 263–275.
52. Kabeya, Y., Mizushima, N., Yamamoto, A., et al. (2004) LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. *J. Cell Sci.* **117** (13), 2805–2812.
53. Kitamura, K., Kishi-Itakura, C., Tsuboi, T., et al. (2012) Autophagy-related Atg8 localizes to the apicoplast of the human malaria parasite Plasmodium falciparum. *PLoS One* **7** (8), e42977.

54. Jayabalasingham, B., Voss, C., Ehrenman, K., et al. (2014) Characterization of the ATG8-conjugation system in 2 Plasmodium species with special focus on the liver stage: possible linkage between the apicoplastic and autophagic systems? *Autophagy* **10** (2), 269–84.

55. Mizushima, N., and Sahani, M.H. (2014) ATG8 localization in apicomplexan parasites: Apicoplast and more? *Autophagy* **10** (9), 1487–1494.

56. Eickel, N., Kaiser, G., Prado, M., et al. (2013) Features of autophagic cell death in Plasmodium liver-stage parasites. *Autophagy* **9** (4), 568–580.

57. Mizushima, N., and Komatsu, M. (2011) Autophagy: Renovation of cells and tissues. *Cell* **147** (4), 728–741.

58. Rai, M.N., Sharma, V., Balusu, S., and Kaur, R. (2015) An essential role for phosphatidylinositol 3-kinase in the inhibition of phagosomal maturation, intracellular survival and virulence in Candida glabrata. *Cell. Microbiol.* **17** (2), 269–287.

59. Yu, Q., Jia, C., Dong, Y., et al. (2015) Candida albicans autophagy, no longer a bystander: Its role in tolerance to ER stress-related antifungal drugs. *Fungal Genet. Biol.* **81**, 238–249.

60. Brennand, A., Rico, E., and Michels, P.A.M. (2012) Autophagy in Trypanosomatids. *Cells* **1** (3), 346–371.

61. Herman, M., Pérez-Morga, D., Schtickzelle, N., and Michels, P.A.M.

(2008) Turnover of glycosomes during life-cycle differentiation of Trypanosoma brucei. *Autophagy* **4** (3), 294–308.

62. Alvarez, V.E., Kosec, G., Sant'Anna, C., et al. (2008) Autophagy is involved in nutritional stress response and differentiation in Trypanosoma cruzi. *J. Biol. Chem.* **283** (6), 3454–3464.
63. Alvarez, V.E., Kosec, G., Sant'Anna, C., et al. (2008) Blocking autophagy to prevent parasite differentiation: A possible new strategy for fighting parasitic infections? *Autophagy* **4** (3), 361–363.
64. Besteiro, S., Williams, R.A.M., Morrison, L.S., et al. (2006) Endosome sorting and autophagy are essential for differentiation and virulence of Leishmania major. *J. Biol. Chem.* **281** (16), 11384–11396.
65. Besteiro, S., Williams, R.A.M., Coombs, G.H., and Mottram, J.C. (2007) Protein turnover and differentiation in Leishmania. *Int. J. Parasitol.* **37** (10), 1063–1075.
66. Williams, R.A.M., Smith, T.K., Cull, B., et al. (2012) ATG5 is essential for ATG8-dependent autophagy and mitochondrial homeostasis in Leishmania major. *PLoS Pathog.* **8** (5), e1002695.
67. Cull, B., Prado Godinho, J.L., Fernandes Rodrigues, J.C., et al. (2015) Glycosome turnover in Leishmania major is mediated by autophagy. *Autophagy* **10** (12), 2143–2157.
68. Meis, J.F., Verhave, J.P., Jap, P.H., and Meuwissen, J.H. (1985) Transformation of sporozoites of Plasmodium berghei into

exoerythrocytic forms in the liver of its mammalian host. *Cell Tissue Res.* **241** (2), 353–360.

69. Jayabalasingham, B., Bano, N., and Coppens, I. (2010) Metamorphosis of the malaria parasite in the liver is associated with organelle clearance. *Cell Res.* **20** (9), 1043–1059.
70. Coppens, I. (2014) Metamorphoses of Malaria: The role of autophagy in parasite differentiation. *Essays Biochem.* **51**, 127–136.
71. Gaviria, D., Paguio, M.F., Turnbull, L.B., et al. (2013) A process similar to autophagy is associated with cytocidal chloroquine resistance in Plasmodium falciparum. *PLoS One* **8** (11), e79059.
72. Vaid, A., Ranjan, R., Smythe, W. a., et al. (2010) PfPI3K, a phosphatidylinositol-3 kinase from Plasmodium falciparum, is exported to the host erythrocyte and is involved in hemoglobin trafficking. *Blood* **115** (12), 2500–2507.
73. Sam-Yellowe, T.Y. (2009) The role of the Maurer's clefts in protein transport in Plasmodium falciparum. *Trends Parasitol.* **25** (6), 277–284.
74. Ghosh, D., Walton, J.L., Roepe, P.D., and Sinai, A.P. (2012) Autophagy is a cell death mechanism in Toxoplasma gondii. *Cell. Microbiol.* **14** (4), 589–607.
75. Sinai, A.P., and Roepe, P.D. (2012) Autophagy in Apicomplexa: a life sustaining death mechanism? *Trends Parasitol.* **28** (9), 358–64.

76. Babbitt, S.E., Altenhofen, L., Cobbold, S.A., et al. (2012) Plasmodium falciparum responds to amino acid starvation by entering into a hibernatory state. *Proc. Natl. Acad. Sci.* **109** (47), E3278–E3287.

77. Engelbrecht, D., and Coetzer, T.L. (2013) Turning up the heat: heat stress induces markers of programmed cell death in Plasmodium falciparum in vitro. *Cell Death Dis.* **4** (12), e971.

78. Totino, P.R.R., Daniel-Ribeiro, C.T., Corte-Real, S., and Ferreira-da-Cruz, M.D.F. (2008) Plasmodium falciparum: Erythrocytic stages die by autophagic-like cell death under drug pressure. *Exp. Parasitol.* **118** (4), 478–486.

79. Lavine, M.D., and Arrizabalaga, G. (2012) Analysis of monensin sensitivity in toxoplasma gondii reveals autophagy as a mechanism for drug induced death. *PLoS One* **7** (7), e42107.

80. Gordon, E.B., Hart, G.T., Tran, T.M., et al. (2015) Inhibiting the Mammalian Target of Rapamycin Blocks the Development of Experimental Cerebral Malaria. *MBio* **6** (3), 1–17.

81. Hain, A.U.P., Weltzer, R.R., Hammond, H., et al. (2012) Structural characterization and inhibition of the Plasmodium Atg8-Atg3 interaction. *J. Struct. Biol.* **180** (3), 551–62.

82. Spangenberg, T., Burrows, J.N., Kowalczyk, P., et al. (2013) The Open Access Malaria Box: A Drug Discovery Catalyst for Neglected Diseases. *PLoS One* **8** (6), e62906.

83. Hain, A.U.P., Bartee, D., Sanders, N.G., et al. (2014) Identification of an Atg8-Atg3 protein-protein interaction inhibitor from the medicines for malaria venture malaria box active in blood and liver stage plasmodium falciparum parasites. *J. Med. Chem.* **57** (11), 4521–4531.

84. Bessoff, K., Spangenberg, T., Foderaro, J.E., et al. (2014) Identification of cryptosporidium parvum active chemical series by repurposing the open access malaria box. *Antimicrob. Agents Chemother.* **58** (5), 2731–2739.

85. Hain, A.U.P., Miller, A.S., Levitskaya, J., and Bosch, J. (2016) Virtual Screening and Experimental Validation Identify Novel Inhibitors of the Plasmodium falciparum Atg8 – Atg3 Protein – Protein Interaction. *ChemMedChem.*

86. Boucher, L.E., and Bosch, J. (2013) Development of a multifunctional tool for drug screening against plasmodial protein-protein interactions via surface plasmon resonance. *J. Mol. Recognit.* **26** (10), 496–500.

87. Ou-Yang, S., Lu, J., Kong, X., et al. (2012) Computational drug discovery. *Acta Pharmacol. Sin.* **33** (9), 1131–1140.

88. Yang, Z., Goronzy, J.J., and Weyand, C.M. (2015) Autophagy in autoimmune disease. *J. Mol. Med.* **93** (7), 707–717.

89. Lavandero, S., Chiong, M., Rothermel, B.A., and Hill, J.A. (2015) Autophagy in cardiovascular biology. *J. Clin. Invest.* **125** (1), 55–64.

90. Menzies, F.M., Fleming, A., and Rubinsztein, D.C. (2015) Compromised autophagy and neurodegenerative diseases. *Nat. Rev. Neurosci.* **16** (6), 345–357.

91. Galluzzi, L., Pietrocola, F., Pedro, J.M.B., et al. (2015) Autophagy in malignant transformation and cancer progression. *EMBO J.* **34** (7), 856–881.

92. Jiang, X., Overholtzer, M., and Thompson, C.B. (2015) Autophagy in cellular metabolism and cancer. *J. Clin. Invest.* **125** (1), 47–54.

93. Boya, P., Gonzalez-Polo, Rosa-Ana Casares, N., Perfettini, J.-L., et al. (2005) Inhibition of Macroautophagy Triggers Apoptosis. *Mol. Cell. Biol.* **25** (3), 1025–1040.

94. Amaravadi, R.K., Yu, D., Lum, J.J., et al. (2007) Autophagy inhibition enhances therapy-induced apoptosis in a Myc-induced model of lymphoma. *J. Clin. Invest.* **117** (2), 326–336.

95. World Health Organization (2015) “Malaria Fact Sheet.”

96. Robert-Gangneux, F., and Dardé, M.L. (2012) Epidemiology of and diagnostic strategies for toxoplasmosis. *Clin. Microbiol. Rev.* **25** (2), 264–296.

97. Huston, C.D., Spangenberg, T., Burrows, J., et al. (2015) A Proposed Target Product Profile and Developmental Cascade for New Cryptosporidiosis Treatments. *PLoS Negl. Trop. Dis.* **9** (10), e0003987.

98. Lévêque, M.F., Berry, L., Cipriano, M.J., et al. (2015) Autophagy-

Related Protein ATG8 Has a Noncanonical Function for Apicoplast Inheritance in Toxoplasma gondii. *MBio* **6** (6), e01446-15.

99. Lionta, E., Spyrou, G., Vassilatis, D.K., and Cournia, Z. (2014) Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* **14** (16), 1923–38.

100. OpenEye Scientific Software Inc. (2012) "OEChem, Version 1.8.1."

101. McGann, M. (2012) FRED and HYBRID docking performance on standardized datasets. *J Comput Aided Mol Des* **26** (8), 897–906.

102. Bennett, T.N., Paguio, M., Gligorijevic, B., et al. (2004) Novel , Rapid , and Inexpensive Cell-Based Quantification of Antimalarial Drug Efficacy Novel , Rapid , and Inexpensive Cell-Based Quantification of Antimalarial Drug Efficacy. *Antimicrob. Agents Chemother.* **48** (5), 1–5.

103. Mizushima, N., Yoshimorim, T., and Levine, B. (2010) Methods in Mammalian Autophagy Research. *Cell* **140** (3), 313–326.

104. Ichimura, Y., Imamura, Y., Emoto, K., et al. (2004) In vivo and in vitro reconstitution of Atg8 conjugation essential for autophagy. *J. Biol. Chem.* **279** (39), 40584–92.

105. Wang, W., Chen, Z., Billiar, T.R., et al. (2013) The carboxyl-terminal amino acids render pro-human LC3B migration similar to lipidated LC3B in SDS-PAGE. *PLoS One* **8** (9), e74222.

106. Talman, A.M., Blagborough, A.M., and Sinden, R.E. (2010) A

Plasmodium falciparum strain expressing GFP throughout the parasite's life-cycle. *PLoS One* **5** (2), 3–7.

107. Yang, Z., and Klionsky, D.J. (2010) Eaten alive: a history of macroautophagy. *Nat. Cell Biol.* **12** (9), 814–822.

108. Klionsky, D.J., and Schulman, B.A. (2014) Dynamic regulation of macroautophagy by distinctive ubiquitin-like proteins. *Nat. Struct. Mol. Biol.* **21** (4), 336–345.

109. Eng, C.P., Sehgal, S.N., and Vézina, C. (1984) Activity of rapamycin (AY-22,989) against transplanted tumors. *J. Antibiot. (Tokyo).* **37** (10), 1231–1237.

110. Powis, G., Bonjouklian, R., Berggren, M.M., et al. (1994) Wortmannin,a Potent and Selective Inhibitor of Phosphatidylinositol-3-kinase. *Cancer Res.* **54**, 2419–2423.

111. Corradetti, M.N., and Guan, K.-L. (2006) Upstream of the mammalian target of rapamycin: do all roads pass through mTOR? *Oncogene* **25** (48), 6347–6360.

112. Yue, Z., Jin, S., Yang, C., et al. (2003) Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. *Proc. Natl. Acad. Sci. U. S. A.* **100** (25), 15077–15082.

113. Kubisch, J., Türei, D., Földvári-Nagy, L., et al. (2013) Complex regulation of autophagy in cancer - Integrated approaches to discover the networks that hold a double-edged sword. *Semin.*

Cancer Biol. **23** (4), 252–261.

114. Reddy, A.S., and Zhang, S. (2013) Polypharmacology: drug discovery for the future. *Expert Rev. Clin. Pharmacol.* **6** (1), 41–7.

115. Peppard, J. V, Rugg, C., Smicker, M., et al. (2014) Identifying Small Molecules which Inhibit Autophagy: a Phenotypic Screen Using Image-Based High-Content Cell Analysis. *Curr. Chem. Genomics Transl. Med.* **8** (Suppl 1), 3–15.

116. Nero, T.L., Morton, C.J., Holien, J.K., et al. (2014) Oncogenic protein interfaces: small molecules, big challenges. *Nat. Rev. Cancer* **14** (4), 248–62.

117. Nevola, L., and Giralt, E. (2015) Modulating protein–protein interactions: the potential of peptides. *Chem. Commun.* **51** (16), 3302–3315.

118. Gasteiger, E., Hoogland, C., Gattiker, A., et al. (2005) Protein identification and analysis tools on the ExPASy server, in *The proteomics protocols handbook*, Springer, pp. 571–607.

119. Hawkins, P.C.D., Skillman, A.G., Warren, G.L., et al. (2010) Conformer generation with OMEGA: Algorithm and validation using high quality structures from the protein databank and cambridge structural database. *J. Chem. Inf. Model.* **50** (4), 572–584.

120. Rasband, W., and U.S. National Institutes of Health Bethesda, Maryland, USA (2014) ImageJ., imagej.nih.gov/ij/.

121. Ma, J., Trop, S., Baer, S., et al. (2013) Dynamics of the Major

Histocompatibility Complex Class I Processing and Presentation Pathway in the Course of Malaria Parasite Development in Human Hepatocytes: Implications for Vaccine Development. *PLoS One* **8** (9), 1–18.

122. McWilliam, H., Li, W., Uludag, M., et al. (2013) Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **41** (Web Server issue), W597-600.
123. The Uniprot Consortium (2014) UniProt: a hub for protein information. *Nucleic Acids Res.* **43** (D1), D204–D212.
124. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42** (W1), 320–324.
125. Sievers, F., Wilm, A., Dineen, D., et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7** (1).
126. Tree Star Inc. (2014) “FlowJo, Version 10.0.6.”
127. The Mathworks Inc. (2014) “MATLAB, Version 8.4.0 (R2014b).”
128. GraphPad Software (2007) “Prism, Version 5.0a.”
129. Persistence of Vision Pty. Ltd. (2004) “Persistence of Vision (TM) Raytracer, Version 3.6.”
130. Schrodinger, LLC (2010) “The PyMOL Molecular Graphics System, Version 1.7.”
131. Cambridge Soft Corporation (2013) “ChemBioDraw Ultra, Version

13.0.2.3020."

132. Idro, R., Marsh, K., John, C.C., and Newton, C.R.J. (2010) Cerebral Malaria; Mechanisms Of Brain Injury And Strategies For Improved Neuro-Cognitive Outcome. *Pediatr. Res.* **68** (4), 267–274.

133. Kyriacou, H.M., Stone, G.N., Challis, R.J., et al. (2006) Differential var gene transcription in Plasmodium falciparum isolates from patients with cerebral malaria compared to hyperparasitaemia. *Mol. Biochem. Parasitol.* **150** (2), 211–218.

134. Leech, J.H., Barnell, J.W., Miller, L.H., and Howard, R.J. (1984) Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. *J. Exp. Med.* **159** (6), 1567–1575.

135. Su, X., Heatwole, V.M., Wertheimer, S.P., et al. (1995) The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. *Cell* **82** (1), 89–100.

136. Bull, P.C., and Abdi, A.I. (2016) The Role of PfEMP1 as targets of naturally acquired immunity to childhood malaria: prospects for a vaccine. *Parasitology* **143**, 171–186.

137. Voss, T.S., Healer, J., Marty, A.J., et al. (2006) A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. *Nature* **439** (7079), 1004–1008.

138. Hain, A.U.P. (2014) "Characterization, structural analysis, and

identification of inhibitors of the Atg8 and the Atg8-Atg3 interaction in the malaria parasite, Plasmodium."

139. Maier, A.G., Rug, M., O'Neill, M.T., et al. (2007) Skeleton-binding protein 1 functions at the parasitophorous vacuole membrane to traffic PfEMP1 to the Plasmodium falciparum-infected erythrocyte surface. *Blood* **109** (3), 1289–1297.

140. Yang, J., Yan, R., Roy, A., et al. (2015) The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **12** (1), 7–8.

141. Zhang, Y. (2007) Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins Struct. Funct. Bioinforma.* **69** (SUPPL.8), 108–117.

142. Zhang, Y. (2009) I-TASSER: Fully automated protein structure prediction in CASP8. *Proteins Struct. Funct. Bioinforma.* **77** (SUPPL.9), 100–113.

143. Xu, D., Zhang, J., Roy, A., and Zhang, Y. (2011) Automated protein structure modeling in CASP9 by I-TASSER pipeline combined with QUARK-based ab initio folding and FG-MD-based structure refinement. *Proteins Struct. Funct. Bioinforma.* **79** (SUPPL.10), 147–160.

144. Zhang, Y. (2014) Interplay of I-TASSER and QUARK for template-based and ab initio protein structure prediction in CASP10. *Proteins Struct. Funct. Bioinforma.* **82** (SUPPL.2), 175–187.

145. Vertrees, J. (2012) "color_by_conservation.py,"

pymolwiki.org/index.php/Color_by_conservation.

146. Korenromp, E.L., Hosseini, M., Newman, R.D., and Cibulskis, R.E. (2013) Progress towards malaria control targets in relation to national malaria programme funding. *Malar. J.* **12**, 18.

147. Murray, C.J.L., Rosenfeld, L.C., Lim, S.S., et al. (2012) Global malaria mortality between 1980 and 2010: A systematic analysis. *Lancet* **379** (9814), 413–431.

148. Mosqueda, J., Olvera-Ramirez, A., Aguilar-Tipacamu, G., and Canto, G.J. (2012) Current advances in detection and treatment of babesiosis. *Curr. Med. Chem.* **19** (10), 1504–1518.

149. Macarron, R., Banks, M.N., Bojanic, D., et al. (2011) Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* **10** (3), 188–195.

150. DiMasi, J.A., Grabowski, H.G., and Hansen, R.W. (2016) Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* **47**, 20–33.

151. Paul, S.M., Mytelka, D.S., Dunwiddie, C.T., et al. (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* **9** (3), 203–214.

152. Scannell, J.W., Blanckley, A., Boldon, H., and Warrington, B. (2012) Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* **11** (3), 191–200.

153. Tan, J.J., Cong, X.J., Hu, L.M., et al. (2010) Therapeutic strategies

underpinning the development of novel techniques for the treatment of HIV infection. *Drug Discov. Today* **15** (5–6), 186–197.

154. Burrows, J.N., van Huijsduijnen, R.H., Möhrle, J.J., et al. (2013) Designing the next generation of medicines for malaria control and eradication. *Malar. J.* **12** (1), 187.
155. Ma, D.-L., Chan, D.S.-H., and Leung, C.-H. (2013) Drug repositioning by structure-based virtual screening. *Chem. Soc. Rev.* **42** (5), 2130–41.
156. Moine, E., Denevault-Sabourin, C., Debierre-Grockiego, F., et al. (2015) A small-molecule cell-based screen led to the identification of biphenylimidazoazines with highly potent and broad-spectrum anti-apicomplexan activity. *Eur. J. Med. Chem.* **89**, 386–400.
157. Bränd´en, C.-I., and Alwyn Jones, T. (1990) Between objectivity and subjectivity. *Nature* **343** (6260), 687–689.
158. Matthews, B.W. (2007) Five retracted structure reports: inverted or incorrect? *Protein Sci.* **16** (6), 1013–1016.
159. Borrell, B. (2009) Fraud rocks protein community. *Nature* **462** (24), 970.
160. Gräslund, S., Nordlund, P., Weigelt, J., et al. (2008) Protein production and purification. *Nat. Methods* **5** (2), 135–46.
161. Janson, J.-C. (ed.) (2011) *Protein Purification: Principles, High Resolution Methods, and Applications*, John Wiley & Sons, Inc., Hoboken, New Jersey.

162. Kabsch, W. (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66** (2), 125–132.

163. Schoeffler, A.J., and Berger, J.M. (2008) DNA topoisomerases : harnessing and constraining energy to govern chromosome topology. **1**, 41–101.

164. Chen, S.H., Chan, N.-L., and Hsieh, T. (2013) New Mechanistic and Functional Insights into DNA Topoisomerases. *Annu. Rev. Biochem.* **82**, 139–170.

165. Liu, L.F., Liu, C.C., and Alberts, B.M. (1980) Type II DNA topoisomerases: Enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell* **19** (3), 697–707.

166. Mizuuchi, K., Fisher, L.M., O'Dea, M.H., and Gellert, M. (1980) DNA gyrase action involves the introduction of transient double-strand breaks into DNA. *Proc. Natl. Acad. Sci. U. S. A.* **77** (4), 1847–1851.

167. Hsieh, T., and Brutlag, D. (1980) "ATP-dependent DNA topoisomerase from D. melanogaster reversibly catenates duplex DNA rings.", *Cell* 21 (**1**), 115–125.

168. Roca, J., and Wang, J.C. (1992) The capture of a DNA double helix by an ATP-dependent protein clamp: A key step in DNA transport by type II DNA topoisomerases. *Cell* **71** (5), 833–840.

169. Roca, J., Berger, J.M., Harrison, S.S., and Wang, J.C. (1996) DNA transport by a type II topoisomerase: direct evidence for a two-gate

mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **93** (9), 4057–4062.

170. Wendorff, T.J., Schmidt, B.H., Heslop, P., et al. (2012) The structure of DNA-bound human topoisomerase II alpha: Conformational mechanisms for coordinating inter-subunit interactions with DNA cleavage. *J. Mol. Biol.* **424** (3–4), 109–124.
171. Nitiss, J.L. (2009) Targeting DNA topoisomerase II in cancer chemotherapy. *Nat. Rev. Cancer* **9** (5), 338–350.
172. Kohanski, M.A., Dwyer, D.J., and Collins, J.J. (2010) How antibiotics kill bacteria: from targets to networks. *Nat. Rev. Microbiol.* **8** (6), 423–435.
173. Tanabe, K., Ikegami, Y., Ishida, R., and Andoh, T. (1991) Inhibition of Topoisomerase II by Antitumor Agents Bis(2,6-dioxopiperazine) Derivatives. *Cancer Res.* **51** (18), 4903–4908.
174. Roca, J., Ishida, R., Berger, J.M., et al. (1994) Antitumor bisdioxopiperazines inhibit yeast DNA topoisomerase II by trapping the enzyme in the form of a closed protein clamp. *Proc. Natl. Acad. Sci. U. S. A.* **91** (5), 1781–5.
175. Fry, A.M., Chresta, C.M., Davies, S.M., et al. (1991) Relationship Between Topoisomerase II Level and Chemosensitivity in Human Tumor-Cell Lines. *Cancer Res.* **51** (24), 6592–6595.
176. Sandri, M.I., Hochhauser, D., Ayton, P., et al. (1996) Differential expression of the topoisomerase II alpha and beta genes in human breast cancers. *Br. J. Cancer* **73** (12), 1518–24.

177. Classen, S., Olland, S., and Berger, J.M. (2003) Structure of the topoisomerase II ATPase region and its mechanism of inhibition by the chemotherapeutic agent ICRF-187. *Proc. Natl. Acad. Sci. U. S. A.* **100** (19), 10629–10634.

178. Hasinoff, B.B., Kuschak, T.I., Yalowich, J.C., and Creighton, A.M. (1995) A QSAR study comparing the cytotoxicity and DNA topoisomerase II inhibitory effects of bisdioxopiperazine analogs of ICRF-187 (dexrazoxane). *Biochem. Pharmacol.* **50** (7), 953–958.

179. Felix, C.A. (1998) Secondary leukemias induced by topoisomerase-targeted drugs. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1400** (1–3), 233–255.

180. Lyu, Y.L., Kerrigan, J.E., Lin, C.P., et al. (2007) Topoisomerase IIβ-mediated DNA double-strand breaks: Implications in doxorubicin cardiotoxicity and prevention by dexrazoxane. *Cancer Res.* **67** (18), 8839–8846.

181. Hawkins, P.C.D., Skillman, A.G., and Nicholls, A. (2007) Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **50** (1), 74–82.

182. Feher, M. (2006) Consensus scoring for protein-ligand interactions. *Drug Discov. Today* **11** (9–10), 421–428.

183. Huang, S.-Y., Grinter, S.Z., and Zou, X. (2010) Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. *Phys. Chem. Chem. Phys.* **12** (40),

12899–12908.

184. Lindsley, J.E. (2001) Use of a Real-Time, Coupled Assay to Measure the ATPase Activity of DNA Topoisomerase II, in *Methods in Molecular Biology: DNA Topoisomerase Protocols*, vol. 95, Humana Press Inc., Totowa, NJ, pp. 57–64.

185. Smith, R. (2011) “doseResponse,” Version 1.2, Retrieved 10/27/15, MATLAB Central File Exchange.

CURRICULUM VITAE

ALEXIA S. MILLER

Born August 15, 1990
In New York, NY, USA

Cell: (302) 388-2730
alexiasmiller@jhu.edu

EDUCATION

Ph.D., Program in Molecular Biophysics, Johns Hopkins University 2012 – 2016

- Estimated GPA: 3.9; 2012-2013 Carlson Fellowship; anticipated graduation June 2016
- Thesis: "Identifying Protein-Protein Interaction Inhibitors Targeting Apicomplexan Atg8-Atg3 Interactions"

B.A., Physics, Princeton University 2008 – 2012

- Graduated *cum laude*; 2012 Kusaka Memorial Prize in Physics; 2011 Kusaka Memorial Prize in Physics; 2009 Academic Competitiveness Grant; 2008 United States Presidential Scholar
- Thesis: "Investigating the Surface Area of *Escherichia Coli* Membranes"

LABORATORY RESEARCH

Johns Hopkins School of Public Health May 2013 – present

Doctoral candidate, Jürgen Bosch's lab, Program in Molecular Biophysics

- Screened, selected, and tested small molecule inhibitors of *Plasmodium falciparum* autophagy-related proteins 3 and 8 and apicomplexan homologues using *in silico*, *in vitro*, and *in vivo* methods

Johns Hopkins School of Medicine October 2014 – present

Graduate student researcher, James Berger's lab, Program in Molecular Biophysics

- Researched and implemented the use of three diverse molecular docking programs (Schrödinger® Glide, OpenEye™ fred, AutoDock™ Vina) for virtual ligand screening to pursue small molecule inhibitors of protein targets of known structure

Johns Hopkins University September 2012 – May 2013

Rotation Student, Program in Molecular Biophysics

- Vince Hilser's lab: wrote Perl™ programs to compare codon compositions of ribosomal genes to the rest of the genome to elucidate selective forces shaping codon composition of ribosomal genes
- Jon Lorsch's lab: developed a fluorescence quenching experiment to monitor the position of eukaryotic pre-initiation complex on a strand of mRNA with respect to the start codon

Princeton University May 2011 – August 2011

Student Research Assistant, Joshua Shaevitz's lab, Physics Department

- To investigate outer membrane tension, I measured the surface area of *E. coli* cell membranes by degrading cell walls and using cytoplasmic and membrane fluorescence to define inner and outer membrane contours, respectively

Princeton University May 2009 – May 2011

Student Research Assistant, Peter Meyers's lab, Physics Department

- Tested photomultiplier tubes and associated electronics, analyzed in MATLAB®
- Modeled and analyzed dark matter detector designs with ANSYS Maxwell®
- Designed and tested prototype capacitive element for measuring changes in liquid Ar

PUBLICATIONS

- **Miller, A.S.,** and Bosch, J. (2016) Targeting the Atg8 Conjugation Pathway for Novel Anti-Apicomplexan Drug Discovery, in *Comprehensive Analysis of Parasite Biology: From Metabolism to Drug Discovery*, 1ed., Wiley-VCH, Weinheim. (Forthcoming)
- **Miller, A.S.*,** Hain, A.U.P.*, Levitskaya, J., and Bosch, J. (2016) Virtual Screening and Experimental Validation Identify Novel Inhibitors of the *Plasmodium falciparum* Atg8 – Atg3 Protein – Protein Interaction. *ChemMedChem*. doi: 10.1002/cmdc.201500515
- Hain, A.U.P., Bartee, D., Sanders, N.G., **Miller, A.S.,** et al. (2014) Identification of an Atg8-Atg3 protein-protein interaction inhibitor from the medicines for malaria venture malaria box active in blood and liver stage *Plasmodium falciparum* parasites. *J. Med. Chem.* **57** (11), 4521–4531.

TEACHING & MENTORING

- **January 2013 – present:** Thread (www.thread.org) volunteer & volunteer coordinator
  - As a volunteer coordinator, supported five mentor-leaders, 15 mentors, and five students, maintaining excellent leadership and student retention
- **Fall 2015:** received certificate from Johns Hopkins Preparing Future Faculty Teaching Academy
- **Fall 2015 & Fall 2013:** Lab Mentor for first year biophysics rotation students
- **Summer 2015:** Co-Instructor for Hands On X-Ray Crystallography Workshop, JHSPH
  - Instructed a class of approximately 15 students during a two-week intensive, interactive workshop for learning how to use a variety of crystallography programs
- **Summer 2014:** Teaching Assistant for Hands On X-Ray Crystallography Workshop, JHSPH
- **Spring 2014:** Teaching Assistant for undergraduate Biophysical Chemistry, JHU
- **July 2009 – May 2012**: Tutor for the Freshman Scholars Institute (FSI) and for the McGraw Center for Teaching and Learning, Princeton University

PRESENTATIONS & POSTERS

- **Presentation:** "Virtually Targeting Apicomplexan Atg8 for Novel Drug Discovery" (December 2015) Student Evening Series, Biophysics Department, Johns Hopkins University, Baltimore, MD, USA
- **Presentation:** "A Virtual Library Screen Hit Inhibiting Malaria Autophagy" (March 2015) Crystal Club Meeting, Biophysics Department, Johns Hopkins School of Medicine, Baltimore, MD, USA
- **Poster**: **Miller, A.S.**, and Bosch, J. (February 2015) "Targeting Apicomplexan Atg8 for Rational Drug Design" Biophysical Society Meeting, Baltimore, MD, USA.
- **Presentation:** "Targeting Apicomplexan Atg8 for Rational Drug Design" (September 2014) Gordon Langsley lab, l'Institut Cochin, Paris, France
- **Poster**: **Miller, A.S.**, Hain, A.U.P., and Bosch, J. (September 2014) "Targeting Apicomplexan Atg8 for Rational Drug Design" Biophysics Departmental Retreat, Baltimore, MD, USA
- **Presentation**: "Truth or Artifact? Controversy in the Cryo-EM of HIV-1 Env" (March 2014) Wednesday Morning Class, Biophysics Department, Johns Hopkins School of Medicine, Baltimore, MD, USA